

- ☐ **Name:** koo do hyun
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

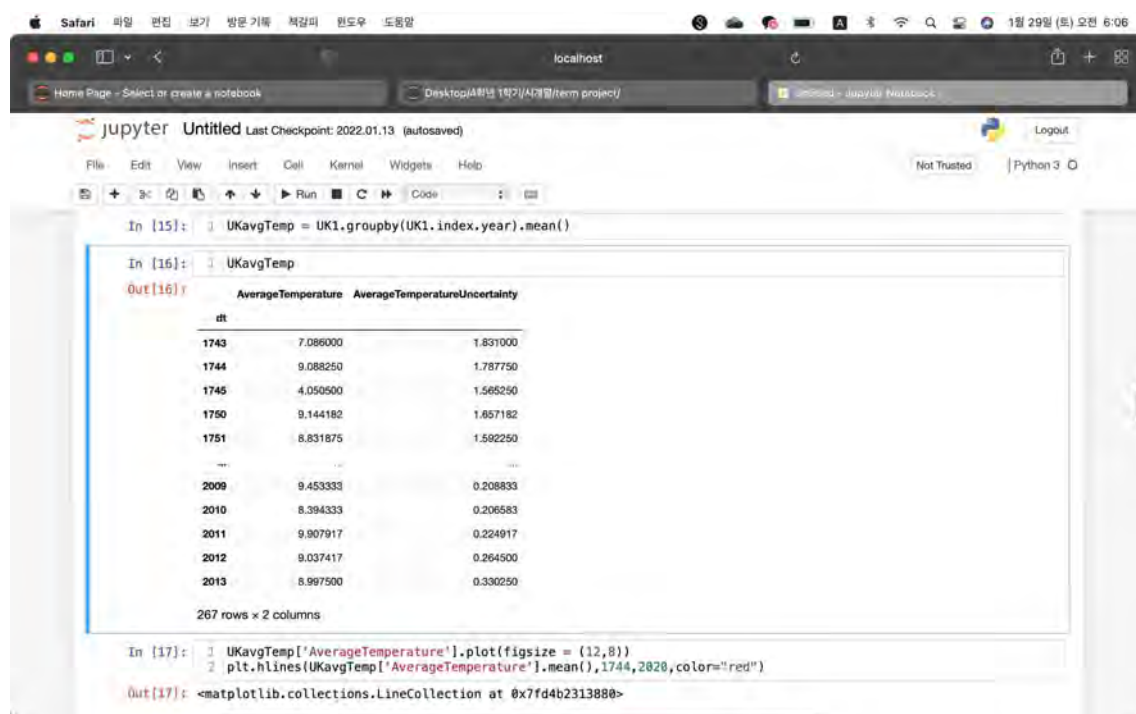
First of all, let's talk about how we feel about the program. My feelings about participating in the program were hopeful, but I want to say it was the worst experience. In a situation where it has become very difficult to go abroad due to COVID-19, I think the purpose of the program was very good to plan and promote such a program to help students experience wider knowledge. In that sense, I was very grateful and happy to give me a hopeful message for widening my horizons by breaking through the blocked path due to COVID-19. However, the reason why the experience finally became the worst is because of COVID-19. If it's because of COVID-19, I think it's because it's not just face-to-face, but it's not for that reason. As a result, it seems that not only the complacent response to both schools and the process of COVID-19 that does not exist before, but also the irresponsible field manager and the irresponsible field manager made my experience the worst. In addition, for the relationship between schools between the two sides, I felt as if I had seen the students as consumables, which was very unpleasant. This is how I feel about hosting the program.

In detail, the contents of the program itself would have been a very interesting class if I had no knowledge in the field. The reason why I said this is that visualizing data is actually a very interesting and difficult task. However, the contents of the class provided by this UIUC were so basic that it was meaningless for me to take the class after studying the department. Therefore, I thought it would be nice to include some level of program content when producing the program. Of course, it is understandable that the project was carried out in a short time. Fortunately, other related classes such as business English, or stories with American companies, can be said to be quite interesting classes.

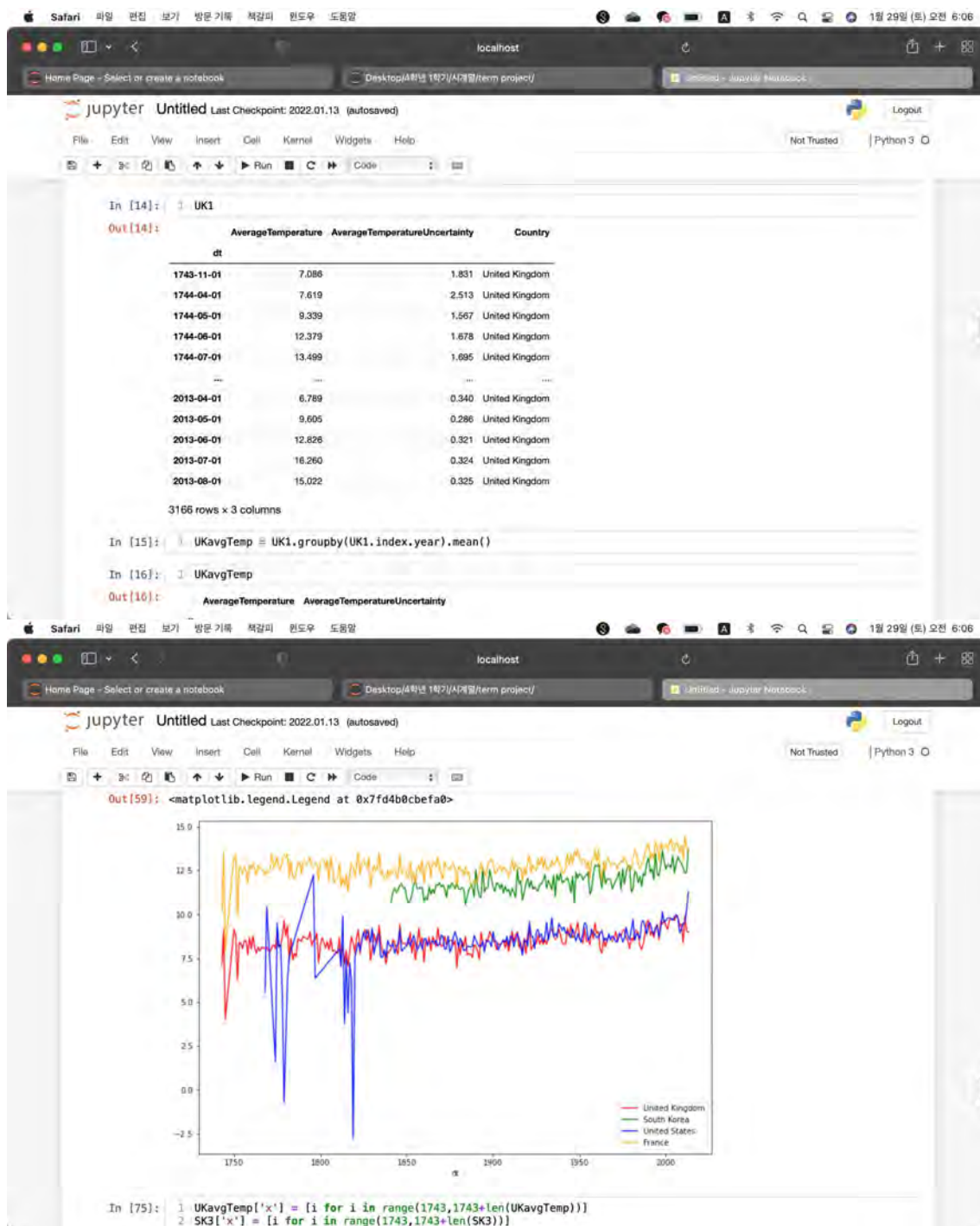
I would like to explain in the report what related projects were carried out this time regarding the class. The project we carried out was curious about how the temperature of the Earth changed. It was to suggest that we should have seriousness at a time when global warming is accelerating. The reason for the content was that there was a threat of global warming, but I wanted to use the power of visualization using the data mentioned by the professor during class. In other words, we can visualize data and become a strong helper to support our opinions as evidence. Basically, after obtaining the data, we loaded the data. After

that, data processing was carried out. Before proceeding with the process, what we thought about was 'How can we show the temperature of the Earth more effectively?' So, the conclusion of our team was to show whether the temperature actually increased despite latitude by setting up countries everywhere to the extreme region based on the equator at the same time as setting up each continent.

Therefore, five countries were selected to show the evidence. Since the data is set as monthly temperatures, we have changed it to annual temperature increases to show an average annual temperature increase, making it easier to see the temperature rise. To tell you the results, the results showed that the temperature of the industrial revolution continued to rise. The data can suggest that continuous temperature rise can pose a major threat to mankind. However, there may be questions about the data truly showing an increase. If you look at the visualization of the data, you can see a sudden drop in temperature it means that it was actually little ice age. Therefore, it was possible to show that the information was actually meaningful through historical information and proof of



facts.



Finishing is a personal opinion on how the program will develop in the future. Now that COVID is becoming more endemic, the basic idea is that there will be more related programs in the future, so that other students can have this opportunity. However, I hope that there will be a more systematic and emergency protocol when conducting the program, so that students will not be anxious when conducting the program, and that the amount of money has been appropriately used. This is because the amount is paid by the people's taxes.

- ☐ **Name: Siun Park**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**

Winter in UIUC

In this winter, I had experienced UNIST Global Program which organized by UIUC in Illinois. When I first arrived in Chicago, the weather was very cold. It snowed a lot, and I can also see the vast land of Illinois. However, there was sun always, the program was well made by UIUC.



The hotel where we stayed was very cozy, clean and comfortable. There was a lot of facilities we needed. We spend most of the time in the hotel room since the COVID-19 situation happened. One of our student was infected to the COVID virus so that we all should do the quarantine in the hotel room. For me, I was also infected to COVID-19, so I did the quarantine alone with listening to the UIUC classes. The classes were well made, but I'm bit disappointed with data science part. I was expected to learn the latest technology and knowledge of data science part. However, what we learn was only the data visualization and some old knowledge of data science. Most of time, we did not learn any of modelling, AI, machine learning techniques. Anyway, the business english

class was my favorite. With casual games, we were excited to join the classes. It was fun, useful and learn a lot of English technique such as small talk, conversation skills, and some of useful words.



This is the picture that I take in the hotel room. The sunrise and sunset of UIUC is amazing. Whole the world becomes amazing color. I often stand at the window and watch the sunset.



This is the place where we take the COVID-19 test in UIUC building. This was the only building that we were able to visit in UIUC. After the day we visit this building, everyone has to be quarantine. This building was used as work out station of UIUC. The facilities were amazing. There are not using now, but they had swimming pool, big court of basketball. I was envy of those station facility levels. The COVID-19 test was done by UIUC's new technique which is using only the saliva to test the COVID-19. The result of test comes really quickly.



After when quarantine was ended (the last day staying in UIUC), finally I was able to visit some campus of UIUC. The university was super large, so I was not able to see all the buildings. These buildings where look like built over 100 years, and it was antique. The downtown of UIUC had everything it needed. It was just a little street with 4~5 blocks. There were several Korean restaurants as well. I tried one of them name "Arirang" and the taste was amazing.



This was the picture taken in last day of UIUC. We were able to see all together at the last day only. Most of time in UIUC, I only have the memory of quarantine.



We went to Chicago downtown after the UIUC schedule. The field trip of Chicago was all cancelled, so we had all the free time for everyone. Those two day was the best day in this program. The downtown of Chicago was awesome, and the I enjoyed this city. I was feeling like having real freedom.

- ☐ **Name:** Kim nanum
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

On January 3, I took the shuttle bus from Chicago O'Hare Airport to the UIUC campus. The accommodation was comfortable. It was enough for two people. I was looking forward to future activities. But sadly, there was a confirmed case of covid19, and all of program was conducted as an online class for two weeks. Some offline activities have been canceled, but classes have been substituted online.

Some classes were too difficult for me as a non-major to understand, but I found them generally interesting. The class I enjoyed the most was the English class. The teachers taught conversation with a variety of activities and topics. I enjoyed listening to it while attending class.

Overall, I was able to hear from various aspects about the fields where AI and big data are used and why they are needed. There were many opportunities to find out what is needed and in what cases it is good to go to UIUC.

We had time to talk with graduate students in online meeting. It was good to hear the stories of experienced people about school life. He also told me about restaurants.

What was disappointing while taking the class was that I felt that it was neither a program for majors nor for non-majors. It was not clear who the target was, so the difficulty of the class was different for each

presenter. I think that non-majors and majors need different lectures, and it would be good if that was clear in the next program.

I prepared for a group project, and although it was an online class, I think I was able to learn a lot. As the program was created with the efforts of many people from preparation to program execution, although it was not the best result, it was nice experience, and I hope that the next participants will enjoy a better opportunity.

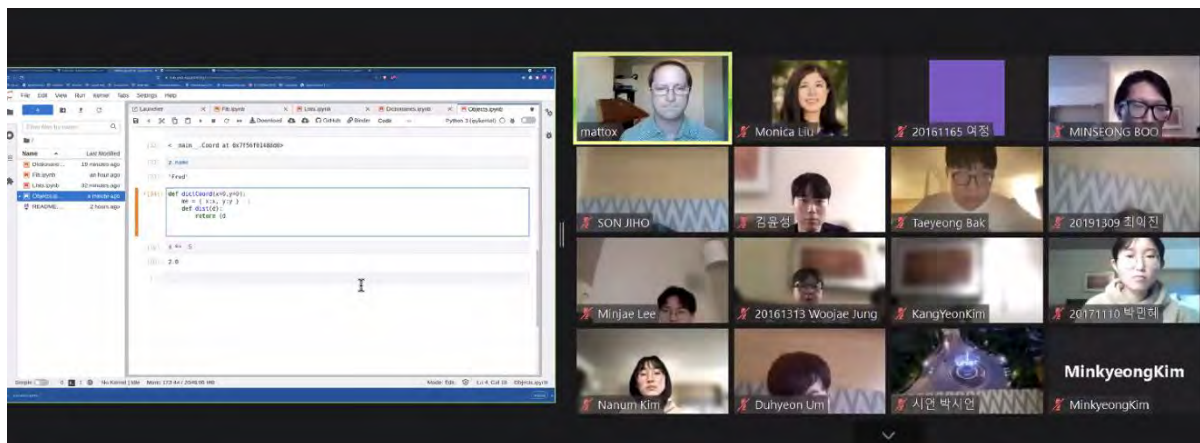
And the personal time given towards the end of the program and the private tour in Chicago were very interesting. It didn't go as planned, but I have no regrets about going.



UGP - UNIST UIUC AI & BIG DATA WINTER CAMP PROGRAM

- ☐ **Name:** YOONSEONG KIM
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

From January 3, 2022 to January 16, 2022, I participated in the UGP program at University of Illinois Urbana-Champaign (UIUC). It was difficult to go abroad for two years because of the corona virus. Therefore, I was excited to be able to go on overseas programs this time. It was also attractive to be able to learn AI and big data, which is the most promising fields in these days. Unfortunately, Bad situations happened one after another and for everyone's health, accordingly, classes were forced to be conducted online.



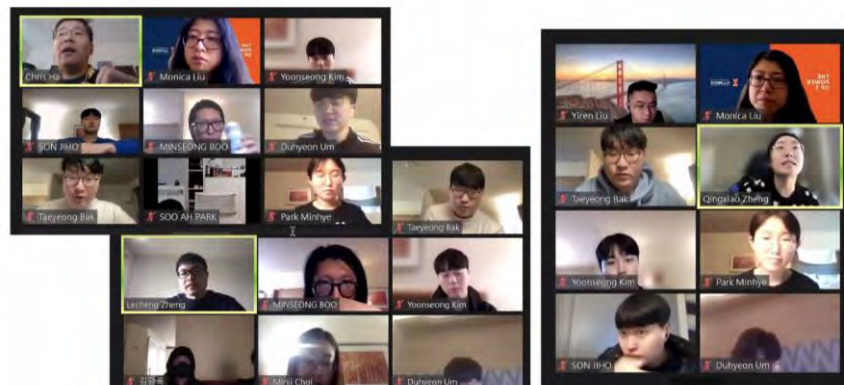
Mattox Beckman, Teaching Assistant Professor: Introduction to Data Structure/Data Visualization/Hands-on Project of Data Analysis

OddBall Algorithm

Proposed method

- For each node,
 - Extract "ego-net" (=1-step neighborhood)
 - Extract features (#edges, total weight, etc.)
 - features that could yield "laws"
 - features fast to compute and interpret
- Characterize regular patterns
 - Examples from **majority classes**
 - Compute the distribution of regular patterns
- Detect irregular patterns
 - Examples from **minority classes**
 - Intuition: far from the fitting line → anomaly

Lecheng Zheng, PhD student in Information Sciences: Machine Learning and Data Mining Models Applied to Social Network Analysis, Healthcare,



Chris Ha, Information and Analysis Manager, Caterpillar Inc. Overview of Enterprise Data Analytics Solutions

HathiTrust origin story

December 2004:
Google & five research libraries announce massive book scanning project.

September 2005:
Authors Guild files a lawsuit against Google and the libraries for "massive copyright infringement."

Glen Worthey, Associate Director for Research Support Services & Ryan Dubnick, Digital Humanities Specialist: HathiTrust Research Center, Hands-on Humanities Data Curation

Thanks to the instructors, I was able to understand a little bit in ai & big data, which was a field I didn't know at all. The course lectures, presentations and projects have met my expectations for introducing new tools and techniques for using and analyzing data. I could see that it was widely used in various fields like agriculture, healthcare, and social network. Through this, I could understand the future industry. In particular, it was an opportunity to feel interested in data analysis. While working on the project, I could feel the importance of tasks such as data cleaning.



Pictures from UIUC campus

Although I couldn't go indoors, the beautiful campus looks wonderful, and I could feel different from Korean universities. In addition, I could feel that the city was formed around the campus. I stayed at a hotel during the camp, but at the end of the camp, I went to Chicago and felt about American culture. Lastly, I would like to thank the school staff because I know that the program would not have proceeded in difficult times without the sincere support of them.

- ☐ **Name:** 부민성
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

After finishing the UIUC program,

During the long and long Corona period, I got an opportunity to participate in the UIUC winter camp as part of an overseas training program supported by the school. Since there was not much time left to leave school, I wanted to participate in various programs, so the Corona situation was even more regrettable. Therefore, I took a risk to participate in the Omikron situation.

Before going to the United States, I was able to arrive at Incheon Airport after a very complicated journey from Jeju Island to Gimpo to Incheon to greet my family. Unfortunately, because the funds were not abundant, I was able to arrive in Chicago after nearly 30 hours of flight through transfer from Doha. I flew four times in three days!

During a short trip in Chicago, I was able to see many things that I had not encountered in Korea. I was able to experience a wide variety of experiences, including trains floating on the city, exhibitions of contemporary art, crosswalk culture, and racism with homeless people. Chicago also had a very lively atmosphere at the end of the year.



After a short trip, I joined the school group going to UIUC and headed to Urbana-Champaign. After being born with the bus for about three hours, I met Monica on Holiday Inn. After hearing a brief explanation from Monica, I was able to rest in a comfortable luxury accommodation.



Unfortunately, there was a COVID-19 infection and I had to start a COVID-19 test and quarantine life in a hurry, and I thought this would end soon, but unfortunately, the journey was carried out until the end.

Based on the class on data visualization, various laboratories, companies, and trends could be introduced. For me, who knew only big data and deep learning used for text processing like Harvest, and deep learning on the image processing side such as chatbot nft, statistical processing, and xai, it was time to broaden my horizons on how and in what direction it can be used in the industry. The stories of people working in various industries such as caterpillar and argo became an opportunity to feel the power of the United States over the economy, experience, and the possibility of attempting.



I would like to thank the staff of the International Cooperation Center for helping me participate in the UIUC Data Engineering program class, Choi Joon-yeop, and UIUC robert, Monica, and teachers who worked hard on the spot.

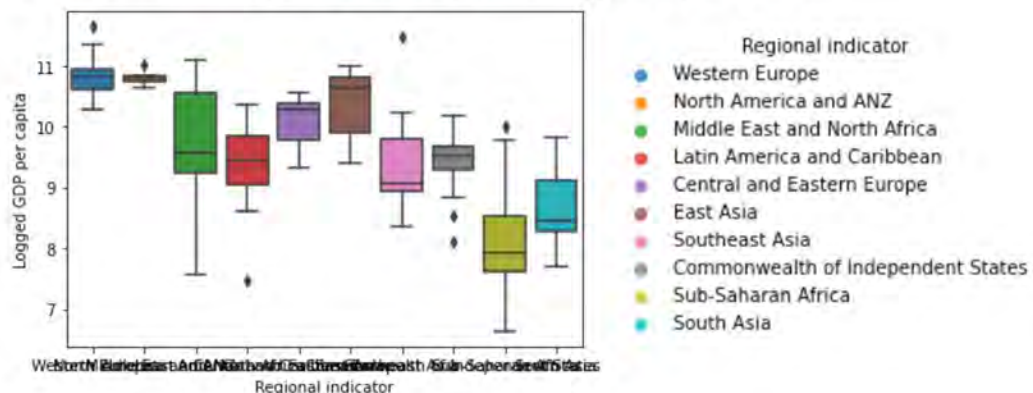
- ☐ Name: 엄두현 / DuHyeon Um
- ☐ Host University/Institute: UNIST
- ☐ Participating Program: UNIST – UGP: UIUC
- ☐ Program Period: 2022-01-03 ~ 2022-01-16

Final Report for UIUC Winter Camp

I participated the UIUC winter camp from 2022.01.03. to 2022.01.16. From 2022.01.02. to 2022.01.03., I departed from Incheon to Chicago. The winter camp was held from 2022.01.03. to 2022.01.13. And from 2022.01.14. to 2022.01.15., I travelled Chicago downtown. Finally, I took a flight to Incheon in 2022.01.16. Unfortunately, this winter camp was conducted by online lecture because of Covid-19 infection of the student. I'm very sad for this decision but it was inevitable. What I was most interested in the lecture is data visualization lecture and practice. In the lecture, I learned the way to make the graph or visualization of dataset to express the trait, trend of the dataset using dataset, python, seaborn program, and pandas program. We practiced it and had a presentation in the last day of the winter camp. We had a presentation using the dataset of the happiness score of the world in Kaggle. In our presentation,

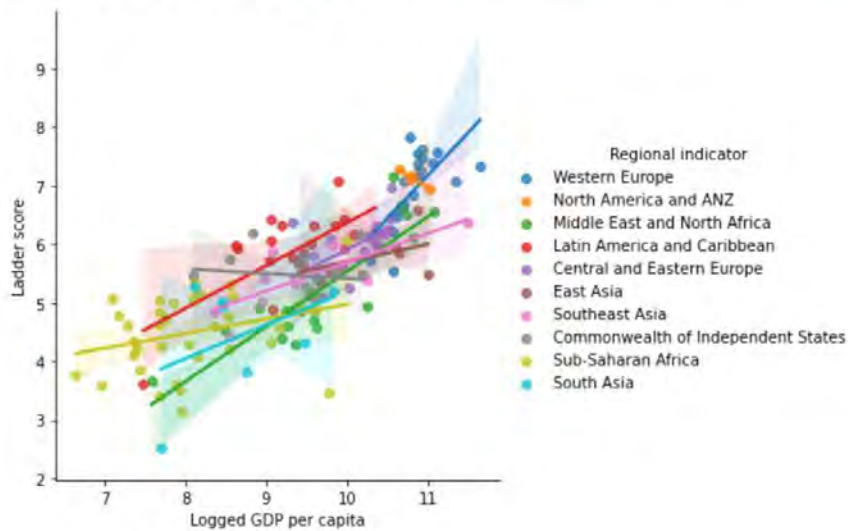
```
[14]: sns.boxplot(data=hap, y = 'Logged GDP per capita', x = 'Regional indicator')
```

```
[14]: <AxesSubplot:xlabel='Regional indicator', ylabel='Logged GDP per capita'>
```




```
[6]: sns.lmplot(data=hap, x="Logged GDP per capita", y="Ladder score", hue="Regional indicator")
```

```
[6]: <seaborn.axisgrid.FacetGrid at 0x2ba0cb4b370>
```



using data visualization, we found the trend or relationship between variables. Even though our group members were beginners of programming, our result was basic practice of data visualization. But it was a great experience to practice the data visualization. Also, There were many lectures from professors, graduate students, presenter from companies, English educators, and personnels of UIUC. Frankly speaking, most of the lectures from the professors, graduate students, presenter from companies are hard to understand for me. The reason is that my major is not related to computer science and data analysis. So, it is hard to understand. But I tried my best to understand the concept which they said at least. By listening the lecture they gave, I was able to be familiar with the speed of the native speakers. By professors and graduates students, I learned about the data cleaning, AI chatbot, machine learning and etc. Especially, the conversation with graduate students of many majors was interested. I was able to talk with them about the life in USA, the mindset of graduate students, studying abroad and etc. From the presenter from the companies, I learned the introduction of the companies and utilization of the data analysis in real field. By English educators, I learned business English like

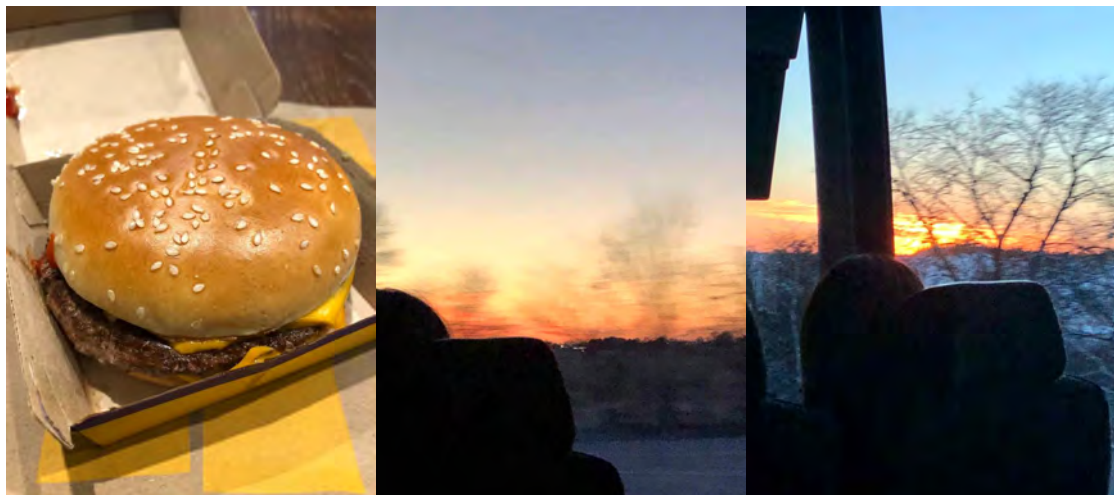
elevator speech and small talk. Last, by personnels of UIUC, I gave the information about the UIUC like I school, admission. The final lecture from prof. Song was interested. He gave a lecture about the AI and future. Also, the lecture was contained the mindset and the traits which companies think importantly in present and future. In the last day of the UIUC winter camp, I walked and saw the campus of UIUC. And during the two days of travel, I saw many things in Chicago downtown like The Art Institute of Chicago, Millenium park, Riverwalk, Willis Tower, and etc. During the two weeks of winter camp, I learned a lot about the data science, AI and big data. Also, I was able to enhance my English listening skill. And, I saw and visited famous places in UIUC campus and Chicago downtown. So, this winter camp will be a good experience in my life.



- ☐ **Name:** Yeo Jeong
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

The overseas training program has finally resumed after two years that I could not even dream of going abroad due to the corona virus. In particular, the US was a place I had never been to before, but I had high expectations for this program because it is a country that I have seen more often than any other country through dramas and movies. This vacation was also the last vacation before going to graduate school, so it was also the last chance to play within the next two years.

We first met together at the airport. At the airport, I saw many Korean-like students, but there was no one who took the lead, so we couldn't get together. There was quite a bit of time left to meet, so I had a McDonald's hamburger for breakfast with a friend at the airport. I didn't know at the time that the hamburger pictured below would become the most delicious food for the remaining two weeks.



We all gathered together and moved to the hotel watching the sunset. And we had a friend who had a pcr test and had a positive result, so we went into quarantine. I was shocked that I was quarantined at the hotel for two weeks because I had seen the friends who had gone to study abroad before the Corona period made foreign friends and continued many exchanges with them.

During quarantine, it made us even more confused, and we were forced to lie down meaninglessly. We took the Zoom class, which is the same as taking an online class in Korea, and even those who got negative results had to stay in the hotel room unconditionally. Fortunately, the hotel room is a bit large, but if it was as small as Korea, someone would have run out saying they couldn't stand it. Here are the foods we were provided.



I felt deeply that American really likes to eat salty at maximum. It was really tasteless, whether it was the restaurant chef's problem or the difference in taste between Korea and America. I just ate because I was really hungry, hoping that every American food wouldn't be like this.

However, once, a staff of UNIST bought a bulgogi ttukbaegi for us so we could eat Korean food. The amount of that food was too much, so I put rest of it in the refrigerator. There was no microwave, so I put it on a heater or tried to heat it by boiling it in hot water. Our plight of having to eat like this was both pitiful and funny.



The classes were conducted well by the professors, but it seems that the difficulty level differed depending on the students because the content they had learned was different. For some it was difficult, for others it felt as a foundation before. I think it would have been more helpful if the classes had been divided into classes according to the levels.

We finished the UIUC program for two weeks while looking at the window view on the right in the hotel room shown in the photo below on the left.



It was a lot of disappointment during UIUC, but it was happy to make new UNIST friends and to make good memories in downtown Chicago.



- ☐ **Name: KANGBEEN LEE**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**



After retiring from army in 2019, I returned to school in 2020 with high expectations, but due to the unprecedented COVID19, classes were held online, and naturally overseas training programs held at UNIST were extended indefinitely. In fact, I was supposed to apply for an overseas training program during the summer vacation of my third year. However, I felt a great regret and gave up my expectations for an overseas training program that had become virtually difficult due to COVID19. However, a few days ago, I saw an announcement on the portal bulletin board that applicants for the 2021 Winter Semester UGP program were recruited. Fortunately, students in the eighth semester can also apply, and as I am in the second semester of the fourth year, I immediately filled out the application form to take advantage of my last chance.

In the case of the UGP (UNIST Global Program), it was a program supported by the UNIST to help students who develop their global capabilities through the AI&Big data project, cultural exchange, and visiting enterprise program. I have been interested in UGP for a long time, and I have worked hard to have OPIC IH level of language skills to participate effectively in UGP. Because of quarantine, every class was proceed in online. So, there was few chance to communicate smoothly with foreign students

and participate in the UGP. As an undergraduate student who wants to go on to artificial intelligence graduate school, I think the AI&Big data course held at UGP is a good opportunity to improve student's ability. Finally, I hope the next program will proceed in offline because I think indoor and outdoor activities and weekend trips in Illinois are also good enough to experience the culture of universities around the world. The below figure shows what our team done in UGP program for project.

```
[87]: six['Region'] = 'NA'
for i in range(len(six)):
    # if six['Region'][i] in region.keys:
    try:
        six['Region'][i] = region[six['Country']][i]]
    except:
        pass

seven['Region'] = 'NA'
for i in range(len(seven)):
    # if six['Region'][i] in region.keys:
    try:
        seven['Region'][i] = region[seven['Country']][i]]
    except:
        pass

eight['Region'] = 'NA'
for i in range(len(eight)):
    # if six['Region'][i] in region.keys:
    try:
        eight['Region'][i] = region[eight['Country']][i]]
    except:
        pass

nine['Region'] = 'NA'
for i in range(len(nine)):
    # if six['Region'][i] in region.keys:
    try:
        nine['Region'][i] = region[nine['Country']][i]]
    except:
        pass

[88]: five['Year'] = 2015
six['Year'] = 2016
seven['Year'] = 2017
eight['Year'] = 2018
nine['Year'] = 2019

[89]: datacolumns = ['Country', 'Region', 'Happiness Rank', 'Happiness Score', 'Economy (GDP per Capita)', 'Health (Life Expectancy)', 'Trust (Government Corruption)', 'Year']
data15 = five[datacolumns]
data16 = six[datacolumns]
data17 = seven[datacolumns]
data18 = eight[datacolumns]
data19 = nine[datacolumns]
```

```
[90]: data = pd.concat([data15,data16,data17,data18,data19])
data
```

```
[91]:
```

	Country	Region	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Health (Life Expectancy)	Trust (Government Corruption)	Year
0	Switzerland	Western Europe	1	7.587	1.39651	0.94143	0.41978	2015
1	Iceland	Western Europe	2	7.561	1.30232	0.94784	0.14145	2015
2	Denmark	Western Europe	3	7.527	1.32540	0.87464	0.48337	2015
3	Norway	Western Europe	4	7.522	1.45900	0.88521	0.36503	2015
4	Canada	North America	5	7.427	1.57629	0.90583	0.32937	2015
...
151	Rwanda	Sub-Saharan Africa	152	3.334	0.35900	0.61400	0.41100	2019
152	Tanzania	Sub-Saharan Africa	153	3.231	0.47000	0.49900	0.14700	2019
153	Afghanistan	Southern Asia	154	3.203	0.35000	0.36100	0.02500	2018
154	Central African Republic	Sub-Saharan Africa	155	3.083	0.02600	0.10500	0.03500	2019
155	South Sudan	NA	156	2.853	0.20600	0.28500	0.09100	2019

782 rows x 9 columns

```
[92]: # 가장 행복한 나라와 가장 불행한 나라
```

```
maxScore = np.max(data['Happiness Score'])
minScore = np.min(data['Happiness Score'])
HPC = data[data['Happiness Score'] == maxScore]['Country']
LMPC = data[data['Happiness Score'] == minScore]['Country']

print('The Happiest Country: ', HPC.values[0])
print('The Most Unhappy Country: ', LMPC.values[0])

The Happiest Country: Finland
The Most Unhappy Country: Central African Republic
```

```
[57]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

import warnings
warnings.filterwarnings(action='ignore')

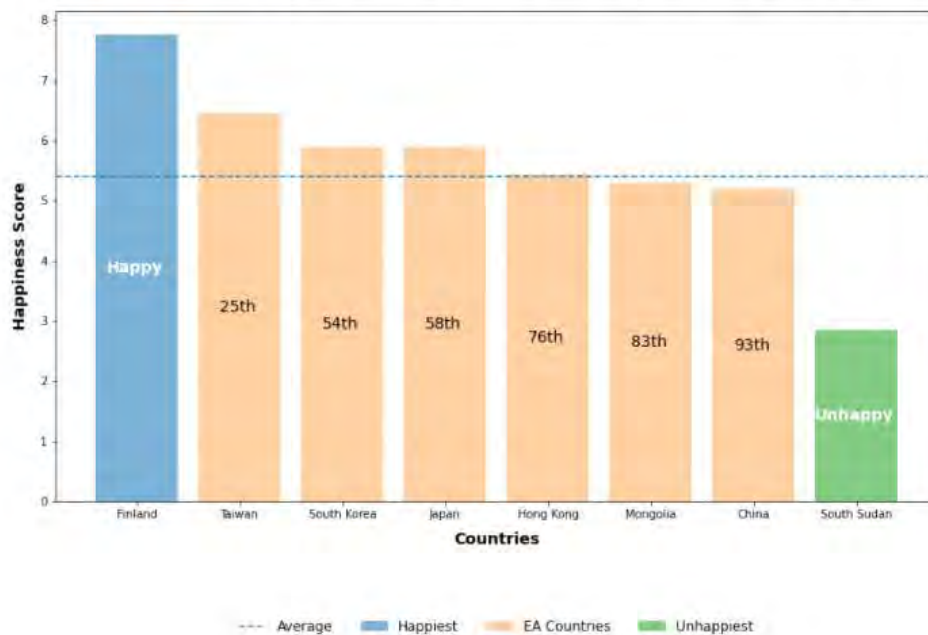
[58]: five = pd.read_csv('2015.csv')
six = pd.read_csv('2016.csv')
seven = pd.read_csv('2017.csv')
eight = pd.read_csv('2018.csv')
nine = pd.read_csv('2019.csv')
df['Stimulation Fluid'] = df['Stimulation Fluid'].str.split(":").str[0]

[59]: seven = seven.rename(columns = {'Economy_GDP.per.Capita.': 'Economy (GDP per Capita)', 'Happiness.Rank': 'Happiness Rank', 'Happiness.Score': 'Happiness Score',
'Health_Life.Expectancy.': 'Health (Life Expectancy)', 'Trust_Government.Corruption.': 'Trust (Government Corruption)'})
eight = eight.rename(columns = {'Overall.rank': 'Happiness Rank', 'Score': 'Happiness Score', 'GDP.per.capita': 'Economy (GDP per Capita)',
'Country.or.region': 'Country', 'Healthy.life.expectancy': 'Health (Life Expectancy)', 'Perceptions.of.corruption': 'Trust (Government Corruption)'})
nine = nine.rename(columns = {'Overall.rank': 'Happiness Rank', 'Score': 'Happiness Score', 'GDP.per.capita': 'Economy (GDP per Capita)',
'Country.or.region': 'Country', 'Healthy.life.expectancy': 'Health (Life Expectancy)', 'Perceptions.of.corruption': 'Trust (Government Corruption)'})

[50]: region_data = five[['Country', 'Region']]
region_data
#region.iloc[:,0]
region = {}
for i in range(len(region_data)):
    region[region_data['Country'][i]] = region_data['Region'][i]

[52]: data_2019 = data[data['Year'] == 2019]
graph(data_2019)
```

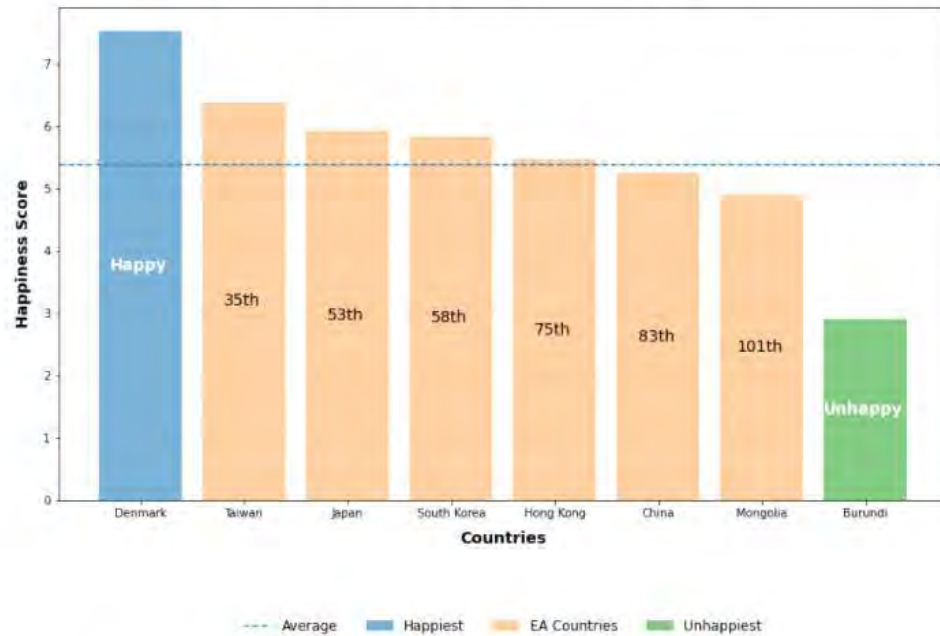
How Happy is South Korea Among its Neighbors?



[]:

```
[50]: data_2018 = data[data['Year'] == 2018]
graph(data_2018)
```

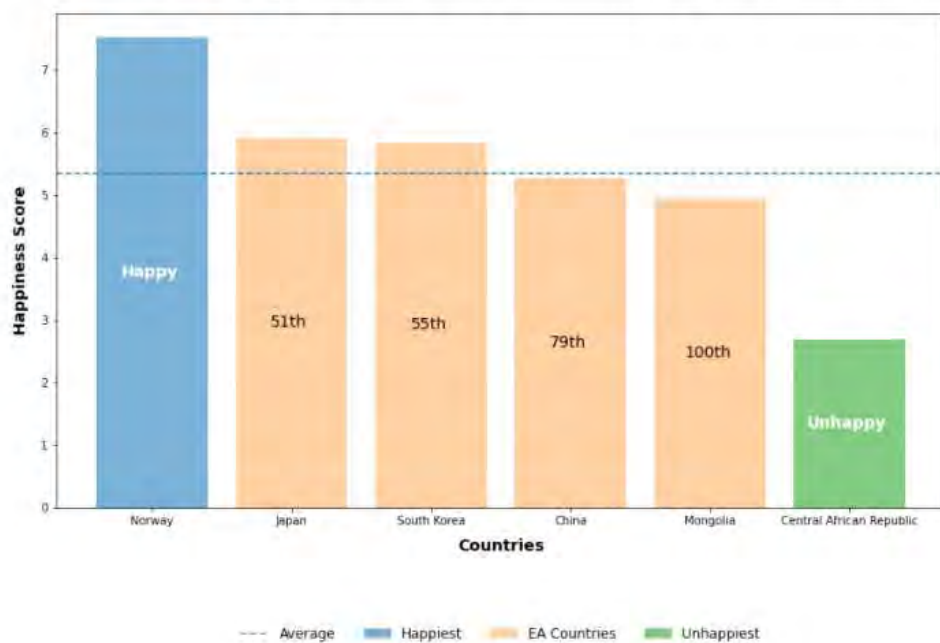
How Happy is South Korea Among its Neighbors?



```
[52]: data_2019 = data[data['Year'] == 2019]
```

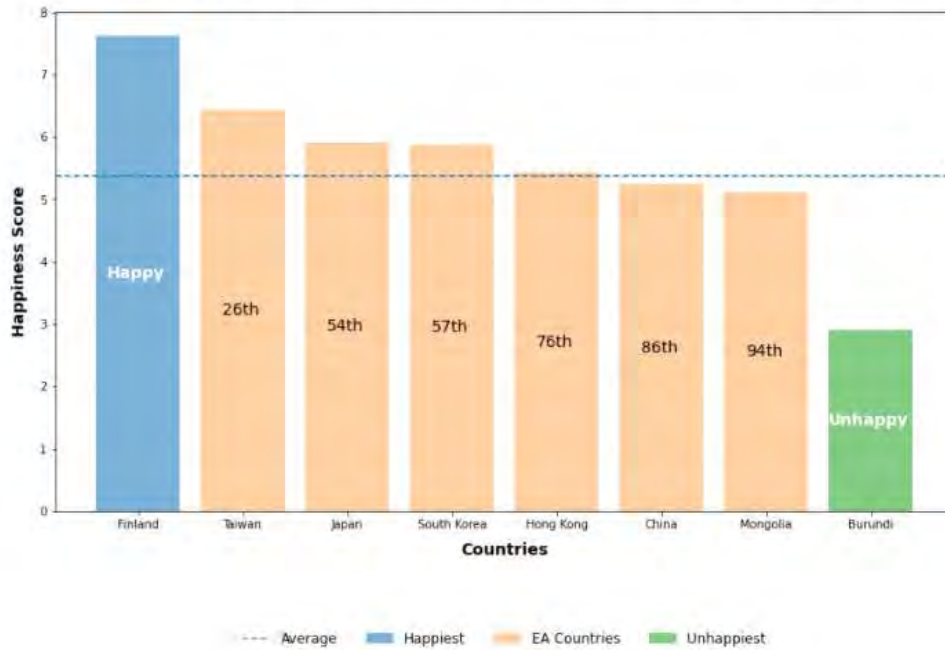
```
[45]: data_2017 = data[data['Year'] == 2017]
graph(data_2017)
```

How Happy is South Korea Among its Neighbors?



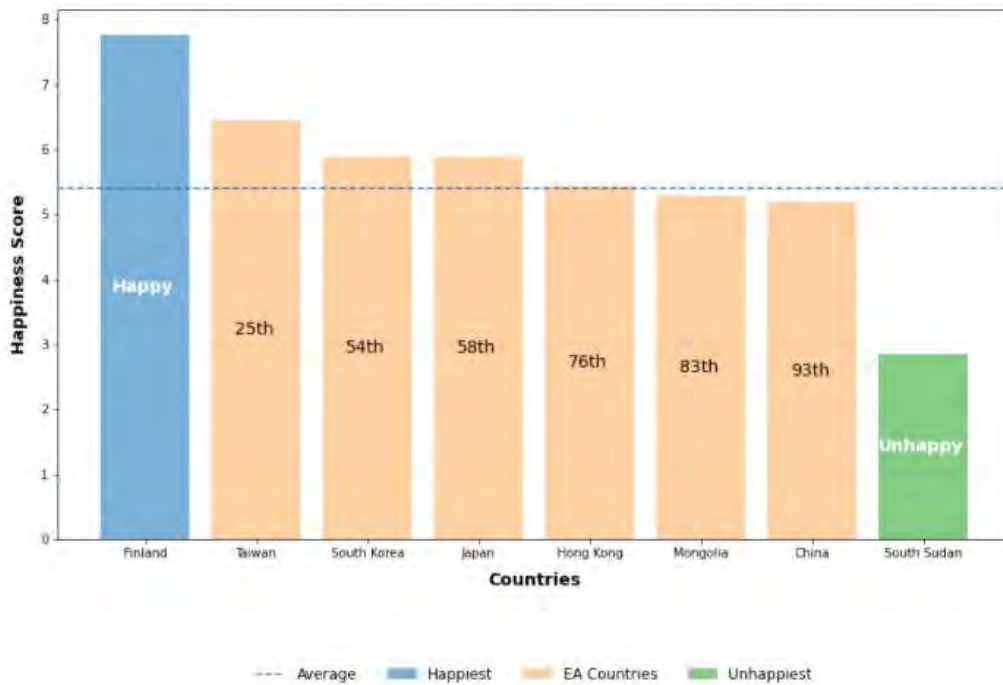
```
[48]: data_2016 = data[data['Year'] == 2016]
graph(data_2016)
```

How Happy is South Korea Among its Neighbors?



```
[49]: data_2015 = data[data['Year'] == 2015]
graph(data_2015)
```

How Happy is South Korea Among its Neighbors?




```

def group(data):
    ea = data[data['Region'] == "East Asia"]['Country'].tolist()
    ea = data[data['Country'].isin(ea)]
    data_top = data.iloc[:1]
    data_bot = data.iloc[-1]
    mean_score = data['Happiness Score'].mean()
    ea_idc = list(ea.index + 1)

    fig, ax = plt.subplots(figsize=(14, 6))

    bars0 = ax.bar(data_top['Country'], data_top['Happiness Score'], alpha=0.5)
    bars1 = ax.bar(data_bot['Country'], data_bot['Happiness Score'], alpha=0.5)
    line = ax.axhline(mean_score, linestyle='--')

    ax.legend(['Average', 'Happiest', 'EA Countries', 'Unhappiest'], loc='upper center', bbox_to_anchor=(0.5, -0.2), ncol=4, borderpad=1, frameon=False, fontsize=12)
    ax.grid(axis='y', alpha=0.3)

    ax.set_xlabel('Country', fontweight='bold', fontstyle='italic', fontweight='bold')
    ax.set_ylabel('Happiness Score', fontweight='bold', fontstyle='italic', fontweight='bold')

    for i, bar in enumerate(bars1):
        xbar = bar.get_x()
        ybar = bar.get_y()
        if i != 7:
            ax.text(
                x=xbar,
                y=ybar,
                text=f'{ea_idc[i]}',
                va='center', ha='center',
                size=12,
                fontweight='bold',
            )

    for i, bar in enumerate(bars0):
        xbar = bar.get_x()
        ybar = bar.get_y()

        ax.text(
            x=xbar,
            y=ybar,
            text=f'{ea_idc[i]}',
            va='center', ha='center',
            size=12,
            fontweight='bold',
            fontstyle='italic',
            fontweight='bold',
            alpha=0.5,
        )

    for i, bar in enumerate(bars2):
        xbar = bar.get_x()
        ybar = bar.get_y()

        ax.text(
            x=xbar,
            y=ybar,
            text=f'{ea_idc[i]}',
            va='center', ha='center',
            size=12,
            fontweight='bold',
            fontstyle='italic',
            fontweight='bold',
            alpha=0.5,
        )

```

```

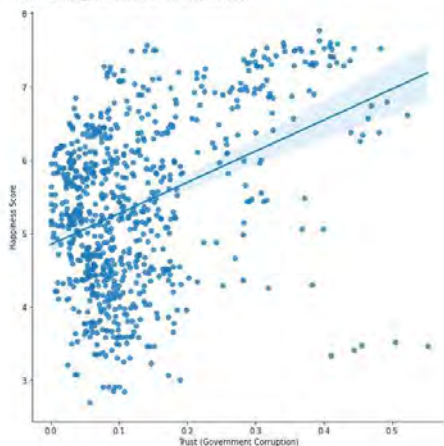
sns.lmplot(x="Trust (Government Corruption)", y="Happiness Score", data=data, palette="summer", height=6)

```

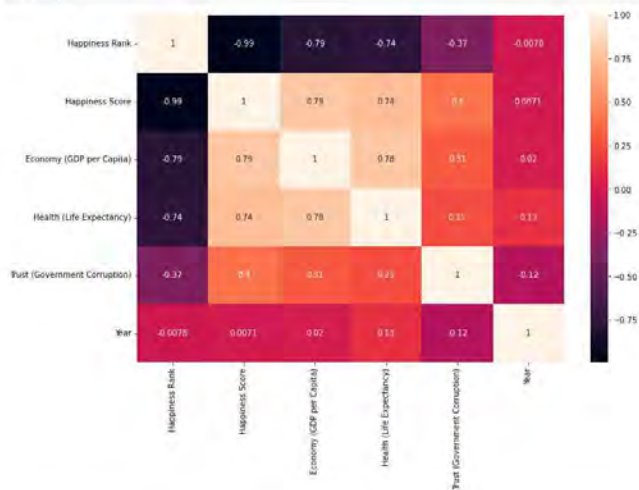
```

<seaborn.axisgrid.FacetGrid at 0x13ee78bb>

```



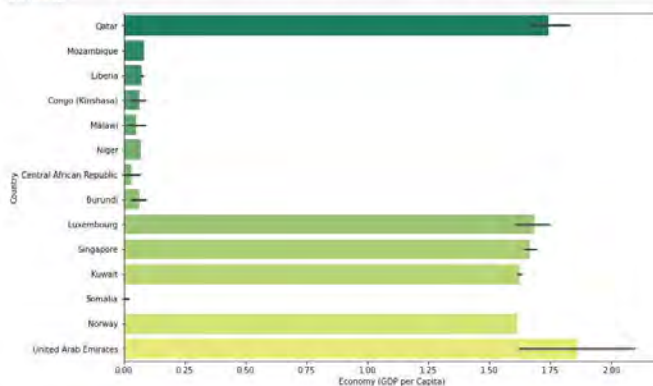

```
plt.figure(figsize=(12,8))
sns.heatmap(data.corr(), annot=True)
plt.show()
```



```
(data["Economy (GDP per Capita)"].describe())
```

```
count    782.000000
mean      0.510047
std       0.407360
min       0.000000
25%      0.000000
50%      0.000000
75%      1.236187
max       2.090000
Name: Economy (GDP per Capita), dtype: float64
```

```
economyFilter = (data.loc[:, "Economy (GDP per Capita)"] > 1.0) | (data.loc[:, "Economy (GDP per Capita)"] < 0.1)
plt.figure(figsize=(12,8))
sns.barplot(x="Economy (GDP per Capita)", y="Country", data=data[economyFilter], palette="summer")
plt.show()
```



UIUC Report

- ☐ **Name:** Minjae Lee
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

형식

- 본 파일 / 최소 3장
- 제목/내용/사진등을 포함한 소감을 적어 자유롭게 작성

UNIST 공식 인정 단계

- 파견 완료 공식 인정 절차
- 보고서 미제출 시, 해외연수장학금
(프로그램 비용, 숙박비, 장학금 차액 UNIST로 반환)

제출 기한

- 프로그램 완료 후, 2주일 이내로 전원 작성 후, PDF 스캔본으로 제출
- 2021.02.13 (일)까지
- 개인 여행 여부와 관계없이 제출

I participated in the UIUC ai course two-week program. The program was conducted as a cyber lecture along with self-isolation because some students had COVID-19. Therefore, it is regrettable that I could not meet professors in person and do not use school facilities. The difficulty level of the ai class was different. However, the last presentation seems to have been a good time for students who are new to ai and students who already have a lot of related knowledge. After self-isolation, I was able to go out for about two days. So I looked around the school. I went into facilities here and there with my friend and took a lot of great pictures. Also, at the end of the program, I had time to spend freely in downtown Chicago. I went to the museum with my roommate and went to the jazz bar while looking at the night view. Also, my high school friend was doing a master's degree nearby, so we met for lunch for a while.

The most memorable thing for me is to walk around uiuc and look around the Chicago downtown. Finally, I will wrap up the article by attaching the photos I took while walking around.















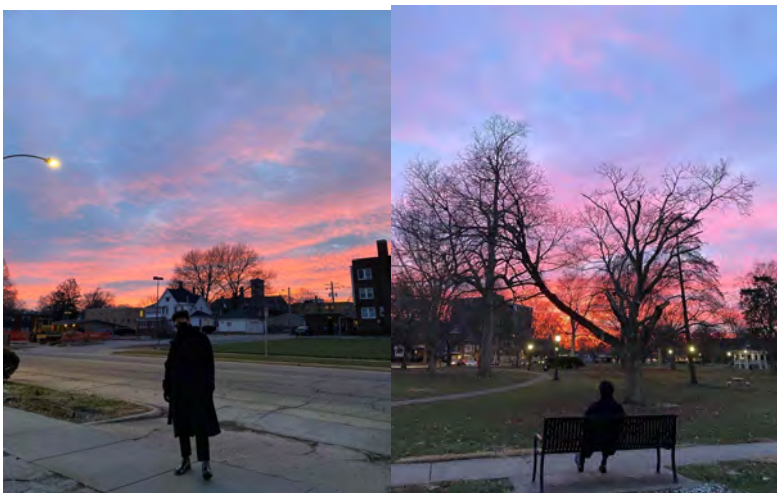
- ☐ **Name:** Yeongbae Jeon
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

AI & Big Data Winter Camp in UIUC

I participated in a program to learn AI and Big Data for two weeks at UIUC in Urbana-Champaign, Illinois, USA. UIUC is a prestigious state university with a focus on research, and I thought it would be a good opportunity to learn about American culture and artificial intelligence.

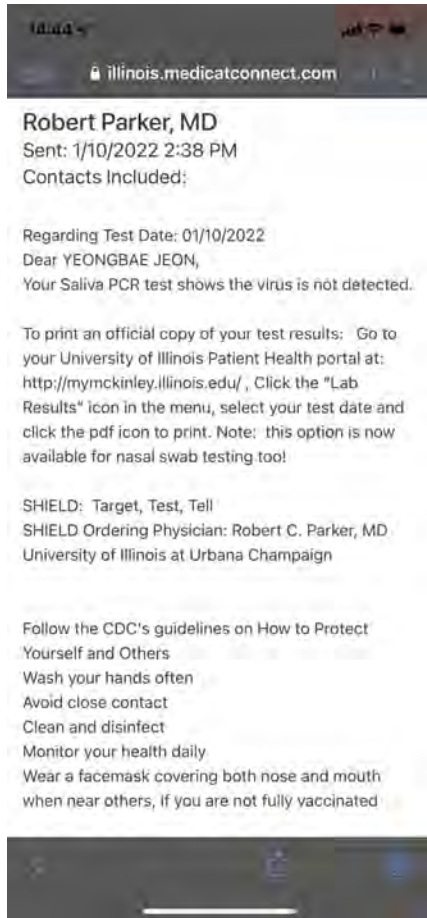


This was my first meal at the Holiday Inn near UIUC. Afterwards, breakfast and dinner were provided by the hostel, so I could learn how Americans eat.



Before being quarantined at the hostel, I went for a walk in the nearby park. There were no tall buildings and the sky was clear, so I was able to enjoy the beautiful sunset.

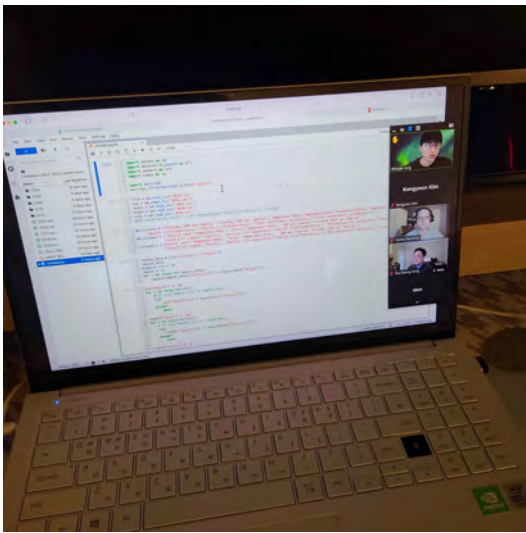
After that, classes were conducted only online through zoom at the hostel until just before the end of the program. It was a pity that I could not experience the campus life at the university in US.



I was able to check the results of the COVID-19 test through the UIUC app. Finally, the quarantine was finished after testing negative for COVID-19. So I took a quick tour of the UIUC campus with my friends.



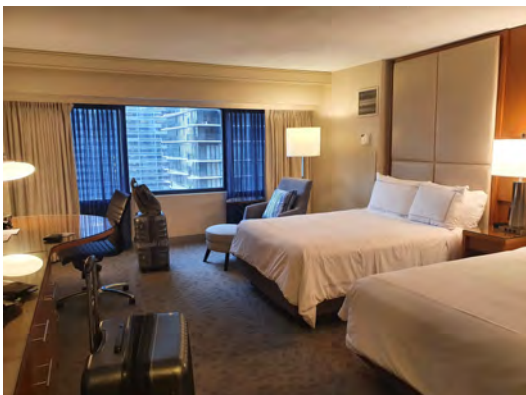
I don't know the name, but I took a picture at the famous statue, and drank a cup of coffee at Starbucks on campus.



At the end of the program, based on what I had learned so far, I coded and presented using data cleaning, analysis, and visualization. Our group held the theme of happiness in countries around the world.



After the program was over, we took a group photo and received a certificate of completion and a souvenir before leaving the dormitory where we stayed for about two weeks.



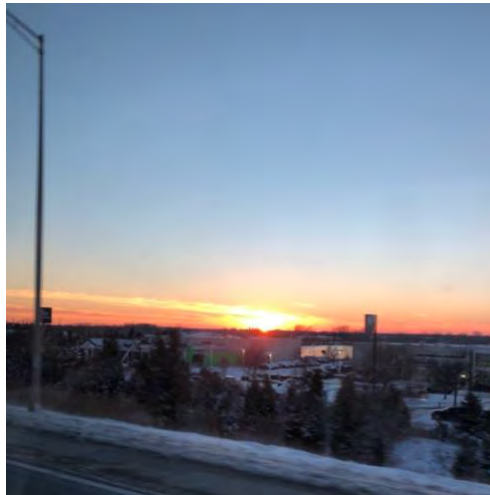
This is where I stayed for 2 nights in Chicago.
It is very unfortunate that UIUC's program was not prepared at all for

COVID-19, but it was beneficial in its own way and I think it can be helpful in my future career.

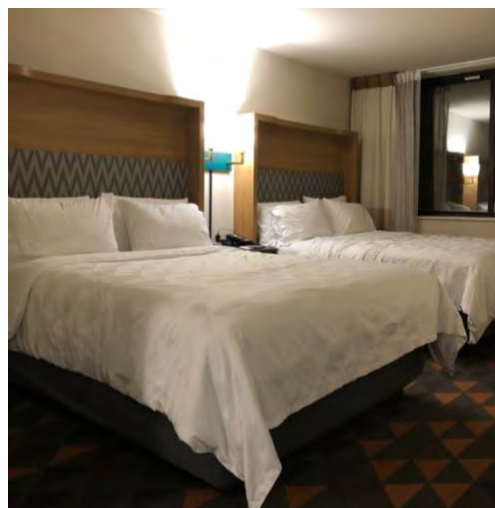
- ☐ **Name: Woojae Jung**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**

Day at Urbana-Champaign

Unfortunately, the program held by the online lecture, the days on Illinois were not forgettable. After meeting the new year on New York city, which is the most memorable day in my life, I made the trip to Chicago through plane at January 3rd. My first image of Chicago was the snow, on the airplane I saw the 'White' airport, that covered with snow. It was much cold than the New York, but it didn't feel that bad because I met my friends who joined to program with me.

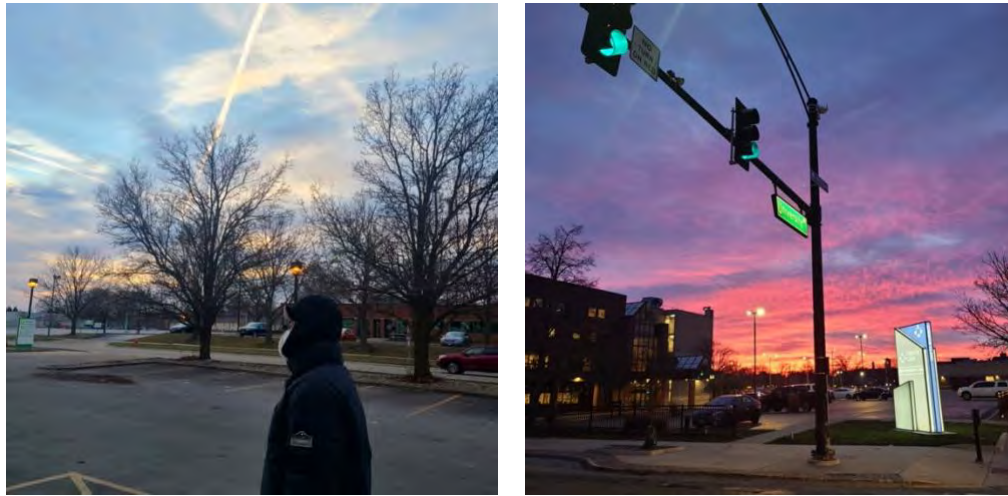


Picture 1, Sight I first saw in the bus to UIUC



Picture 2, Holiday Inn, I lived for two weeks

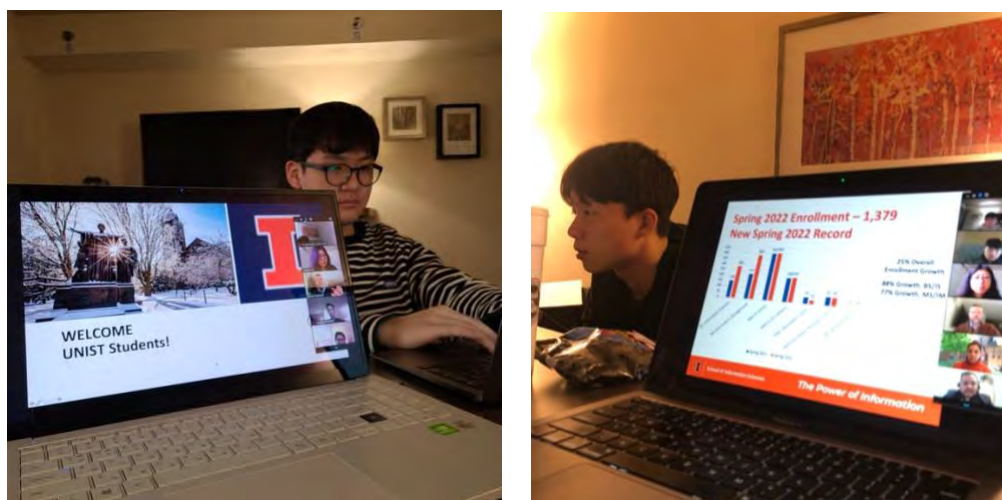
Initial program plan was the three days of online lecture and visited the UIUC and get lecture in the offline. So, I made the plan to visit the Urbana-Champaign with my roommate, 강연; I already knew he is the fancy and nice man, I knew he is but also clean and cool by this time.



Picture 3, Beautiful sky at UIUC at day and sunset

Lecture of UIUC

To be honest, it was not the program I thought. I imagined that learning about the programming related to machine learning. For example, doing the project of text mining using the python or JAVA whatever the developer tools or theoretical lecture about the machine learning by the famous professor in UIUC. Sadly, the half of the lectures were held by the Ph.D. candidate student, where was really disappointing. But that did not mean that I hate all the lectures. Lecture by Mattox who was our core lecturer teaching us about programming.



Picture 4, I and Kangyun listening to Lecture

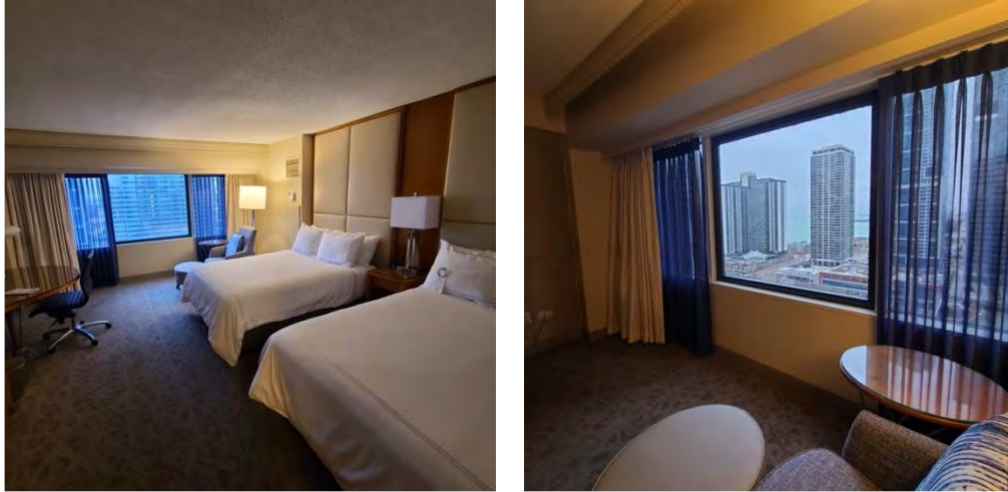
The Business English lecture was funny and good, but it was bit easy to be learned on the abroad of my hometown. For the last project, I team up with my friends and we made a discussion to select the best data to show what we've learned in lecture. It was the data that Mattox had failed in the class as the data was not well organized as the data was written in different format in each year.



Picture 5, Discussion time in UIUC Starbucks with my team!!!

Day at Chicago

Not only for me, but majority of the students might also be most excited of the Chicago downtown tour in this program. Luckily most of the students tested in negative at covid-test. We got a chance to tour the Chicago. First, we were happy to live in this fantastic hotel by the help of the UIUC. Chicago was beautiful city as I thought. Visiting The Art Institute of Chicago was great experience to saw the memorable art works. Also, the food was tasty, especially deep-dish style pizza and the buffalo wing! Only three days in Chicago made me to dream about living in this beautiful city.



Picture 6, Fantastic hotel I lived for multiple days in Chicago



Picture 7, Happy time in Chicago

Conclusion

This program did not be held like the plan, but it was truly great experience for me. The lecture can't be said that was great, but it was meaningful to know there were many different areas related to the machine learning. The great part of this program was that I can feel the different culture of USA and by living lecture of English to tour the city, I can be more confident in my listening and speaking skills.

□ Name: 갈민지

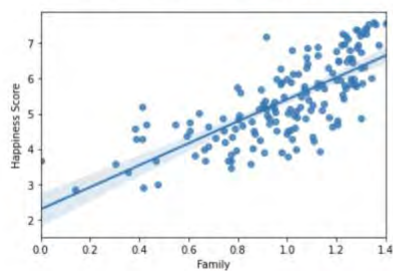
□ Host University/Institute: UNIST

□ Participating Program: UNIST – UGP: UIUC

□ Program Period: 2021-01-03 ~ 2021-01-16

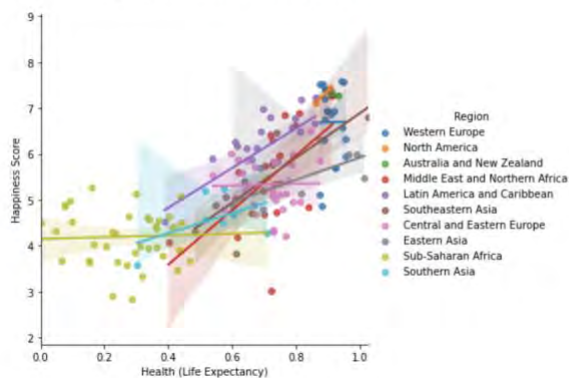
```
In [5]: sns.regplot(data=happy, x="Family", y="Happiness Score")
```

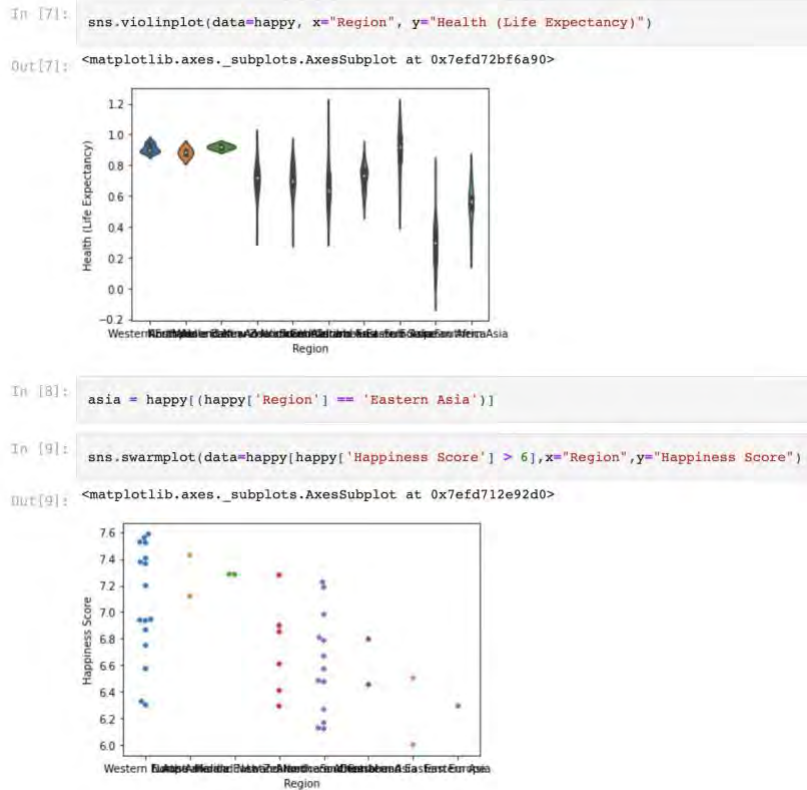
```
Out[5]: <matplotlib.axes._subplots.AxesSubplot at 0x7efd75436bd0>
```



```
In [6]: sns.lmplot(data=happy, x="Health (Life Expectancy)", y="Happiness Score", hue="Region")
```

```
Out[6]: <seaborn.axisgrid.FacetGrid at 0x7efd72cd9550>
```





교수님이 제공해주신 예시코드.

영어, 빅데이터와 인공지능 분야의 다양한 경험과 산업소개 등의 수업들이 진행됐는데, 메인으로 진행되던 파이썬을 이용해서 데이터를 시각화하는 수업에서 배운 것을 마지막날 발표를 했다. 특별한 아이디어나 영감을 얻을 수 있는 수업들은 아니었지만 평소에 흔히 사용하지 않는 형태의 그래프를 직접 그려봤다.



침을 검체로 이용한 pcr 검사를 받아보는 새로운 경험을 했다.



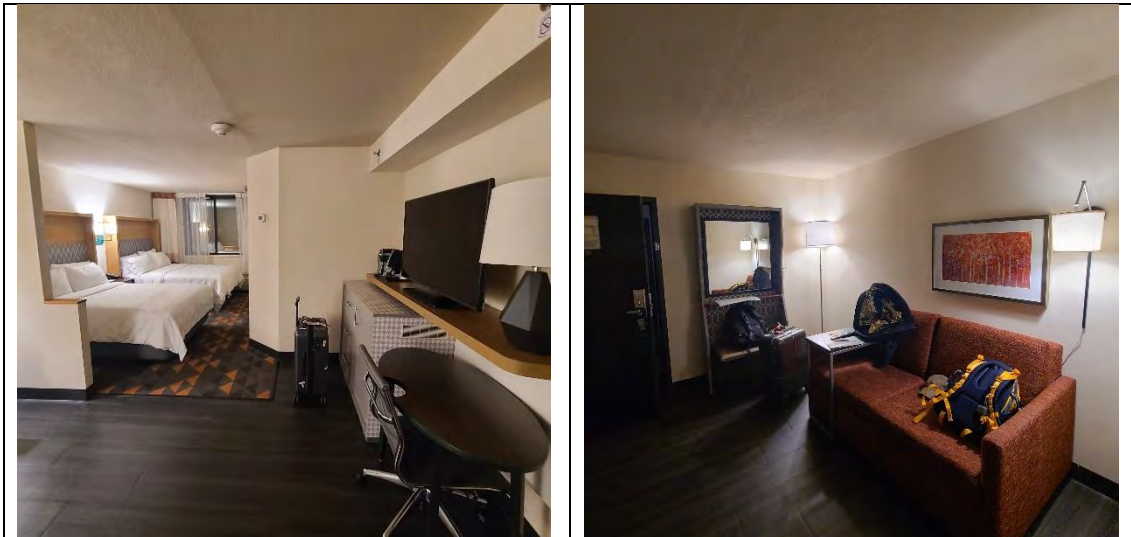
거의 모든 시간을 호텔 방에 갇혀서 지냈으며, 식사 또한 도시락으로 제공되었다. 한국인지, 귀한 돈과 시간을 들여서 온 미국인지 느끼기 어려웠으나, 우리나라의 대부분 실내공간과 달리 어둡고 색온도가 낮은 조명, 입맛에 맞지 않는 식사들과 함께하는 것으로 해외에 있는 기분을 느낄 수 있었다.



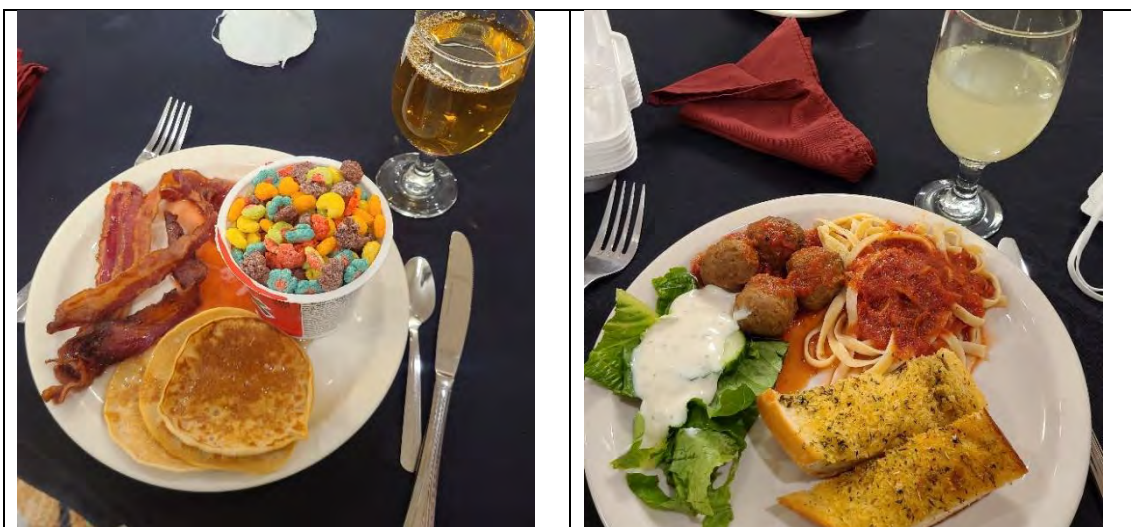
마지막날 찍은 단체사진. 첫날에 잠시 마주치고, 노트북 화면 너머로만 보던 학우들의

얼굴을 직접 볼 수 있었다.

- ☐ **Name:** KangYeon Kim
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

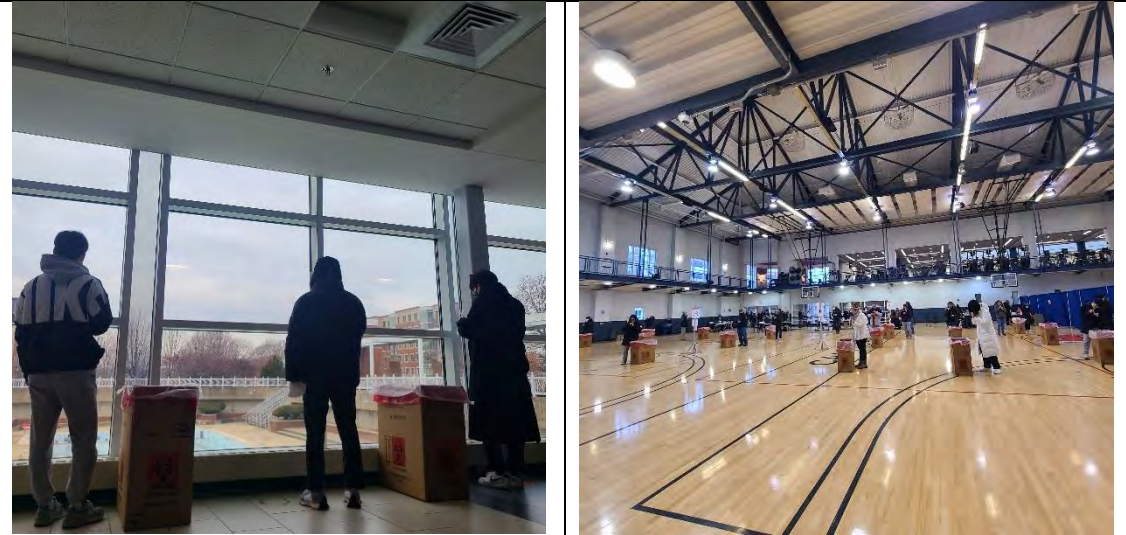


When I arrived at O'Hare Airport after a long flight, I was surprised by the weather in Illinois which is colder than Korea. I ate McDonald's as first meal in the United States. After a hard waiting, I finally got on the bus going to hotel. When I arrived at the hotel, I quickly unpacked my luggage and washed. It felt like the difficult time of uncomfortable flight and waiting time for two days was melted away.



Breakfast and dinner were served at the hotel. Eggs, hamburgers, bacon, pancakes,

cereals, and more meals were served which I think as typical American meals. Lunch was usually eaten at a restaurant near the hotel. It was delicious, but honestly, it reminded me a lot of Korean food.

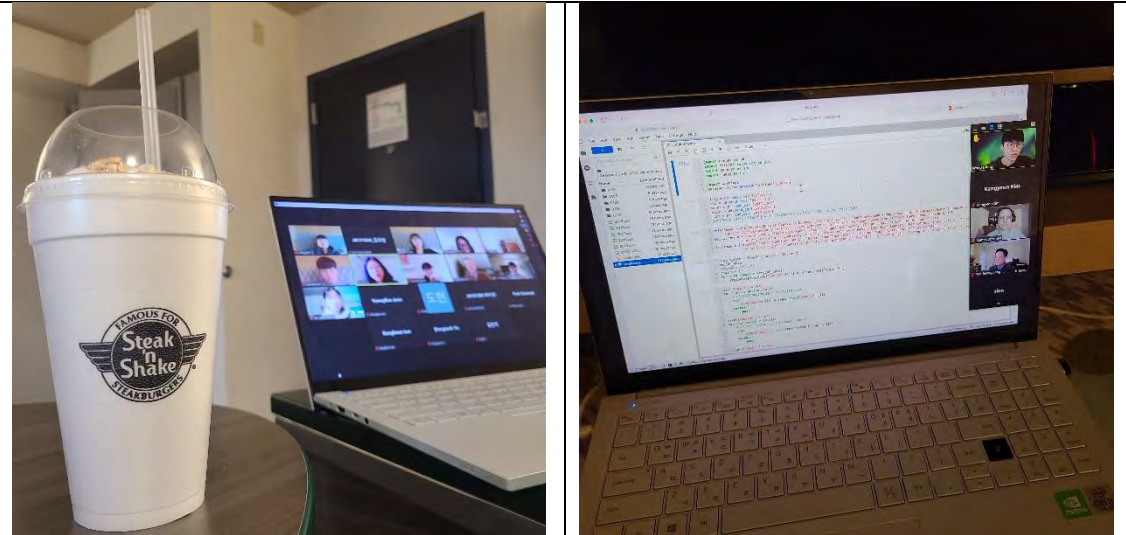


I was tested for covid-19 several times while participating in the UGP program. It was called the 'Saliva PCR Test' which was test covid through saliva. There were several positive results Among those who participated in the program. Thus, I got in quarantine and all session are replaced by online classes. It was unfortunate, but I felt fortunate that no one was seriously ill.



During the quarantine period, I was able to get lunch through a platform called 'Uber Eats'. I mainly ate Korean food by this platform. The taste was slightly different

from that of Korea, but it was very delicious. Also, I ate pizza delivered to a group, and I was impressed with the taste.



Through the online class, I listened to lectures related to the AI field, graduate schools, and companies. On the last day, some group members performed presentation about their data project.



After time in Champagne, I moved to Chicago. In this new place, I was able to enjoy my free time on the program. I liked new hotel shown in picture very much, and I was visited many tourist attractions and famous place in Chicago. Only except for the cold weather in Chicago, it was my best time in UGP program.

- **Name: Minkyong Kim**
- **Host University/Institute: UNIST**
- **Participating Program: UNIST – UGP: UIUC**
- **Program Period: 2021-01-03 ~ 2021-01-16**

1. Visualize data.

I learned the concept of data visualization and proceeded with the project.

The seaborn module for data visualization and the kaggle site for easy data collection were found.

2. panel discussion

It is said to be one of the common techniques for improving design.

I knew it for the first time as I didn't major in design.

It was interesting because anyone could give an opinion and anyone could participate without paying attention.

3. Meeting with international students.

I was able to ask international students about their study fees and preparation process.

It was a good opportunity.

4. Corporate briefing.

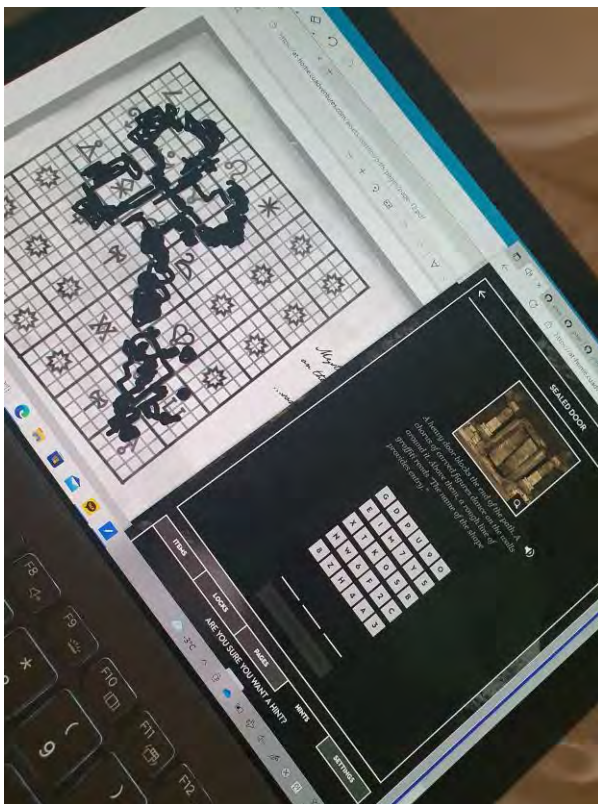
To be honest, I don't know why I did the briefing. It was embarrassing that we, not even international students, suddenly heard the explanation of foreign companies.

5. Business English.

It was interesting to learn about small talk.

It was good to learn the courtesy of the English culture.

6. Zoom class.



To be honest, I didn't remember anything and lost interest in taking the zoom class.

It was a question class that was far from enough to write a speech.

UIUC did not show an attitude to improve this quality class.



The quarantine was lifted only for about two days, so I had time to look around the school.

7. Inn

The accommodation itself was not bad, but the food served was very bad.

They didn't provide a proper diet.

8. A free trip to Chicago.

The accommodation was good.

Traveling was a bit difficult because the weather was cold,

I could feel the feeling of traveling abroad for a while.



- ☐ **Name:** 박민혜 Park Min Hye
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

New country, exciting, a lot of people, different people, diversity, COVID-19, sick, guilty, sorry, quarantine, zoom again, disappointed, some inspiration, motivated, thanks home back, and finally I realized let's refrain from unnecessary overseas tour when it is pandemic.

As a person who wants to be an onsite disaster expert after my study, I expected to learn following things from this program; 1) Data generation and management skills even under the vulnerable situations (as most disaster-prone areas are as so), 2) Communication skills with various people who have different backgrounds and history from me (as I might face a lot of them during my career), 3) Being out of my comfort zone (to increase my survivability). Reminding them again after the program finished, I could achieve most of them.

1) Enhancing data generation and management skills

It was very unfortunate to have all the lectures online which makes me feel this program as no more than any other online AI courses that can be easily found in Korea as well. It was quite different experience from what I expected, what we all expected, but still there were some wonderful lecturers and I could also learn a very useful free online tool to trim the raw data (<https://openrefine.org/>). Professor Mattox introduced himself who has a Korean wife, he was very kind and caring person and he provided us many useful code lines to plot various types of graph. I learned some new plotting code lines including violet plot and I think those can be practically used on my study or even during my whole career. Especially, while working on the final presentation, I was very motivated and the process itself made me think about a lot of technical, ethical problems that I have to consider when I manage data. Students like me who are interest on some practical research that can bring some direct effects on our real life, we cannot talk about the 'research' without 'data'. Through the data, I see the world, recognize a problem, then I can do some good research that might solve the problem. So, it is all about data and if we go further, its visualization as well. I really enjoyed his classes and learned a lot of data visualization skills. Another professor, Song Yoo-Seong, his lecture was inspiring and motivated me. He reminded me that it is our role to decide what to say from the data. People are going to believe what we deliver them. Data can be interpreted in many different ways by many different researchers. I feel this like we are a reporter who delivers daily news to the people. The people obviously will be affected by everything delivered from the reporter and I

believe I should be an honest reporter, honest researcher who can carry the information with clear objectivity and accuracy and be a nice storyteller. To be the one, I am going to keep studying on understanding the different point of views, diversity, and the nature of data.

2) Better communication skills with various people

I love to learn different people. I value all different lives. It was so beautiful to see the diversity in the US. We say the US is a melting pot and that is the most wonderful thing when I think about the US. Reminding my days back in the US, still and maybe even stronger, I feel the diversity of the people is the most amazing thing in the US. No one would feel like an alien in the US. Maybe New York city the most as I visited when the city was especially full of diverse people which was Christmas season. Even though some Chinese speaking strangers talked to me in Chinese, I did not mind them, rather happy. There were yellow, black, white, queer, slim, fat, Muslim, christian, vegetarian, ... I could see all the human diversity there. I had a few chances to talk to them and what I realized was just we are all same human. For the first few days, I felt so fresh and brand new walking on the street, but 3 days later, all became so familiar and even felt like I am just enjoying my vacation in Korea. Just if I am ready to accept a little distance from the other people and willing to jump into their daily experiences, communicating with strangers will not be a matter. Also, while I stay in the US, I really thought a lot about my international friends in Korea. All the awkward and uncomfortable situations that they had to went through, all the eyes toward them, and all the unstable and not-belonging feelings. It must be very tough for my friends especially in Korea where is mono-ethnicity society. So, I think I could also learn the better communication skills with my international friends in Korea.

3) Being out of my comfort zone

I cannot talk about my journey in the US without COVID-19 infection. It is the main of the whole story. I never had a single cough last year, so it was a quite shock to me to get confirmed. I could not plan my journey accordingly and it brought me a big fear at first. Even worse, as it was a powerful infectious disease, I was a risk to other people around me. Rather than feeling physical pain, I was suffered by feeling guilty. All the happenings caused by COVID-19 infection put me totally out of my comfort zone and I was trained harsh. Fortunately, I could come back home on the scheduled day, but still it gave me a lot of teaching. I had to think about many solutions in case I cannot go back to Korea and I believe I could learn a little bit of 'Flexibility of life' through the experience. I do not want to learn it again in that way, but I feel it helped me to grow. Reminding myself let's do not plan unnecessary travels especially abroad when it is pandemic, I would like to close this writing.



Figure 1 Day



Figure 2 and night in Chicago



Figure 3 Trex



Figure 4 UIUC campus



Figure 5 UIUC campus 2



Figure 6 Dinner

- ☐ **Name:** Hyowon Ki
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

형식

- 본 파일 / 최소 3장
- 제목/내용/사진등을 포함한 소감을 적어 자유롭게 작성

UNIST 공식 인정 단계

- 파견 완료 공식 인정 절차
- 보고서 미제출 시, 해외연수장학금
(프로그램 비용, 숙박비, 장학금 차액 UNIST로 반환)

제출 기한

- 프로그램 완료 후 2주일 이내로 전원 작성 후, PDF 스캔본으로 제출
- 2021.02.13 (일)까지
- 개인 여행 여부와 관계없이 제출

UNIST Global Program at UIUC

I participated in the UNIST Global Program (UGP) held from January 3 to January 16 at the University of Illinois Urbana-Champaign (UIUC). UGP consisted of various lectures and projects on AI, data science, machine learning, and informatics. I took lectures such as Data Visualization, Hands-on Humanities Data Curation, Introduction to Research Park and Enterprise Works Incubator, Data Analytics and IoT and Interdisciplinary Research at Illinois. I had a memorable time for two weeks, even though all the schedules were switched to non-face-to-face due to COVID-19, and all lectures were provided using ZOOM.

I worked on a team project in a Data Visualization class by a professor Beckman Mattox at Grainger College of Engineering. I learned how to analyze and visualize data and practiced codes to draw graphs using Python. Finally, I carried out a team project that drew meaningful or interesting conclusions after visualizing and analyzing the data. Our team selected data from individual passengers on Titanic. The data includes gender, age, survival, parch, number of siblings, fares, and cabin classes. Among them, we wanted to see how the proportion of survivors varies by age and cabin class. We visualized the data by drawing violin graphs and bar graphs using the codes we learned in class. As a result, it was concluded that the older the age and the higher the cabin class, the

higher the probability of survival. Through this hands-on project, I was able to learn how to visualize vast amounts of data and use various codes. In addition, there were several teams that chose Titanic data, and it was also interesting that the analysis and interpretation of the same data were different.

Since I am a senior, I had a lot of worries about my future career. I am planning to enter UNIST graduate school after graduation, but I was also curious about studying and researching at graduate schools abroad. Participating in UGP helped me think about and decide on my career after graduation. Because I could have time to talk directly with the graduate student of UIUC. In addition, I took lectures on current careers in AI and Data Science. I was able to ask questions and get useful answers about things I was usually curious about. Through this program, I was able to get a lot of information that could help me make career decisions and learn about the advantages and disadvantages of attending graduate schools abroad and how to prepare for admission.

In addition to lectures dealing with academic content, we took business English classes. I was able to learn about the rule and etiquette of conversation in the United States. This class was interesting to me because I was only used to listening to English or reading textbooks and papers, and I was not confident in conversation. In addition, learning about conversational etiquette was good because it was an opportunity to learn about the culture of the country.

The scheduled campus tour was canceled due to COVID-19, but I conducted a self-campus tour with my friends in our free time. UIUC has a very large campus and the largest library system after Harvard. We looked around the Granger Engineering Library, the largest engineering library in the United States. The spacious space, high ceilings, and free atmosphere were impressive. Most of the students were studying while discussing with their friends, not alone. Also, on the last day, I traveled to Chicago. I made unforgettable memories while traveling to Chicago with my friends and wrapped up my two-week schedule.

I didn't get to take classes on campus because all schedules were

changed non-face-to-face due to COVID-19 and quarantine. However, this program was meaningful to me because I could listen to good lectures on AI and data science, meet new friends, and make unforgettable memories.



1. The campus



2. The Granger Engineering Library



3. The Granger Engineering Library



4. Travel to Chicago

- ☐ **Name:** 손지호, SON JIHO
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

UGP report

Purpose and Object of participation

From 1/3~1/16 I participated UIUC illinois campus program. I once lived in united state to follow my father who work as professor. I residue at the professor apartment. The memory of united state at that point was really great. The people's attitude, atmosphere, and culture gave me a great impact that I always thought that I should go back to america if I have a chance. I thought this program was the chance to satisfy my desire and opportunity to use the overseas training scholarship I earned for hardwork at highschool.

Activity



In the day we arrived at the university of illinois, we went straightly to our hotel. We met Monica at here who was one of the person in charge of the program. Since we needed time to adjust to the time difference we rested at hotel room

First week activity was proceeded like the below image. We had lots of schedule to use the school facility and meet school facility representative in person.

1p Introduction to Data Structure	8a Covid Testing	9a Hands-on Humanities Data Curation	9a Responsible Computing with Social Data: Bias Detection, Theory Validation, and Impact Assessment	10a THE LOST TEMPLE Virtual Game!
3p GET Staff Checking In and Q&A	9:30a Machine Learning and Data Mining Models Applied to Social Network Analysis, Healthcare, Financial Fraud Analysis	1p Introduction to Business English	11a Professor Beckman Mattox' Office Hour	11:30a JET'S Pizza Delivery
	1p Introduction to iSchool Enrollment Management, Student Affairs, Academic Affairs	3p Introduction to Research Park and Enterprise Works Incubator	1p Business English: Presentation Tips and Tools I	1p Data Projects
	3p Data Visualization		3p Enterprise Data Analytics Solutions	

The first thing we did at Illinois was to get COVID-19 test at school gym to get the building access. Unfortunately, one student in our group had a positive result at the test causing all the first week program to be an online class.



The content of the first lecture was how machine learning and data acquisition can be applied to school's financial, health, network and so on. To be honest, this lecture was hard for me to understand. However, the next lecture which covered basic Python functions was easy for me.

The next day's topic was about HathiTrust Research Center, business English, entrepreneurship incubator. Each presenter told us about the topic for 1.5~2 hours.



Hands-on Humanities Data Curation with the HathiTrust Research Center

Glen Worthey

Associate Director for Research Support Services

Ryan Dubnick

Digital Humanities Specialist

HathiTrust Research Center

School of Information Sciences

The University of Illinois Urbana-Champaign

HathiTrust is an inter-cooperation group of non-profit academic research libraries which reserves more than 17 million digitalized data. And it provides access to the maximum amount of data allowed by U.S. copyright law and allows computerized access necessary for academic research. In this organization, some services and programs are being conducted to develop missions and goals, of which the HathiTrust Research Center provides services to analyze HathiTrust's corpus as a dataset through data mining or text mining research.

In the lecture of the Business English we learned some words, expressions, proverbs and phrases that we might use frequently when we do business at U.S. We divided into teams and practiced these words and phrases to each other. This

lecture was covered for two weeks which I thought it was most helpful to me since I was planning to get a job earlier than others after I graduate UNIST.



In the third lecture of day, I was able to know that at lots of start-up enterprise study at the illinois research park. This illinois research park had lots of equipment, lab and environment for the start-up team to develop well. UNIST also has many start-up teams, but it was regrettable that there was no such environment and on the other hand, I was very envious.

Next day lectures were composed of Responsible Computing with Social Data, Business English.

Responsible Computing with Social Data was a research that was hold in illinois data lab which accumulates human's social behavior data to solve the social issues in daily life. I was interested from the fact that a social issues and problems can be expressed in the form of data and solved by data science.

After the weekend we had another Corona 19 test. We were supposed to have an offline class if we got negative at the test, but the plan was canceled which meant we had to have the whole class at the hotel room.

9	10	11	12	13	14	15
12:00a Provide Data Project Group List	6a IMPORTANT Covid Testing	9a Data Cleaning and Analysis to Support Reproducible Research Projects	9a Machine learning explainability and its applications	9a Camp Student Showcase/Presen	6a Covid Testing	Chicago on your own
1p Data Projects	9a Hands-on Project of Data Analysis	11:15a Interaction and Panel Discussion with IS/Engineering Graduate Students	1p Program Updates with Matt, Rob, and Monica	11:45a [Tentative] Online Recognition Ceremony	9a [Tentative] The state-of-art Engineering, Manufacturing, and AI at MxD	
	1p Ph.D. Student Panel Discussion on Conversational AI	1p Business English: Presentation Tips and Tools II	2p AI and Big Data careers	3p Where do we go from here? ~ careers in AI and data science	12p Depart for Chicago	
	3p Data Analytics and IoT	3p Interdisciplinary Research at Illinois	3p Business English: Presentation Tips and Tools III			

Last week of the program was scheduled like above.

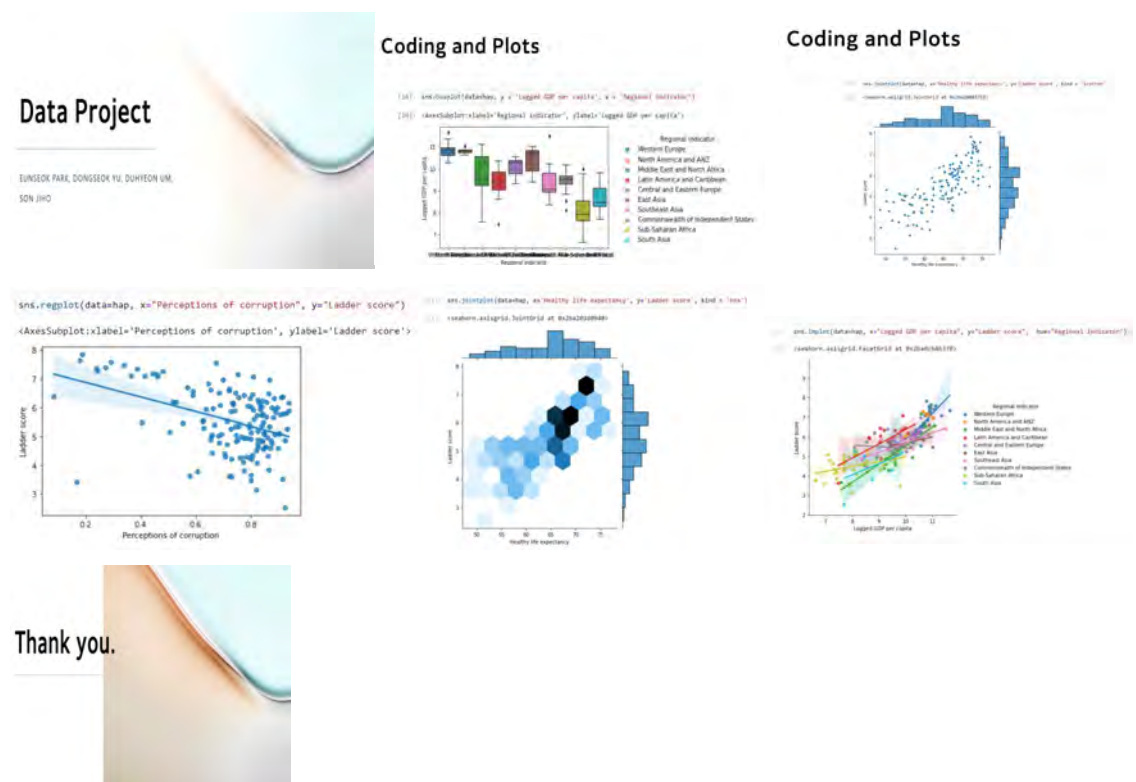
In Monday class we prepared the final presentation that will be take place on Thursday. We were told to visualize the data using seaborn, jupyterlab that we learned at the first week. Also in Tuesday class we learned the mechanism to clear and augment the datasets which we needed for the presentation. This class was really helpful for our presentation team because our team had an difficulty choosing the datasets.

In addition there was a lecture which was held by graduated students in illinois. They told us about the benefit, environment, finacial part of school life and contact method. Since it wasn't easy to have a contact with the students in illinois graduate student in daily life, it was imformative lecture for me.

Besides of these lecture, few company manager presented about the company's vision, future, internship, target and so on.

All of the companies were new to me and the area they apply AI and data science was pretty surprising. The company who use AI to collect the error of machines such thing as lithium battery, excavator and change them into dataform to provide an solution to the useres were most memorble company. Company who mines and programs data to provide an way to increase efficiency in agricultural area was also interesting to me.

Finally Thursday was the day we had presentation.



Our group used 2021 happiness datasets from kaggles. We visualize graph of the

datasets using seaborn and jupyterlab to find the correlation between variant such as GDP, corruption, happiness and so on. Using these correlation we presented what we found in these data and what matched our expectations.

After the presentation we had Online Recognition Ceremony by zoom.



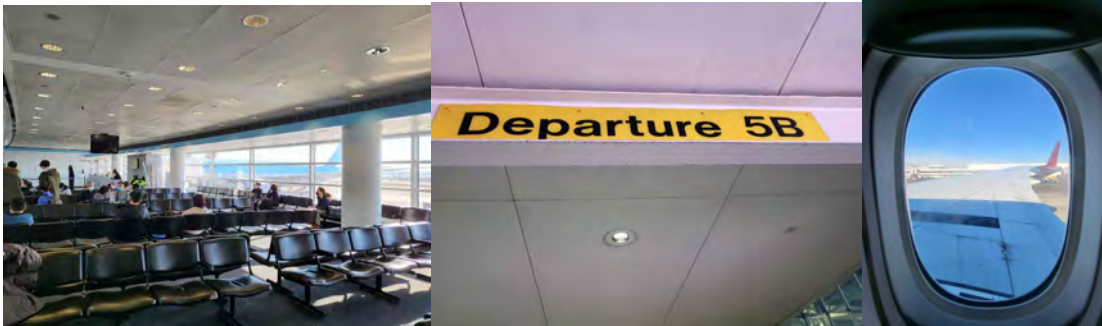
After we finished all the courses, before going for Chicago we took pictures in front of the hotel, received certification and illinois uniform. It was too bad that this was the only time we met with group in this program.



In the plan, we were supposed to visit MxD engineering, Manufacturing, AI industry applications facility and have Chicago style pizza as a dinner in Chicago. However, the plan was totally revised, resulting us to travel by our own in both Friday and Saturday. So I traveled with Duhyeon Um who was my roommate.



In Sunday 6 am we went to O'hare airport made our way back home.



Review for the program

The good side of this program was that we had presentations from variety of people in the university. Since those presenters used AI and data in all different parts in the university, I was able to know the large area that AI can affect.

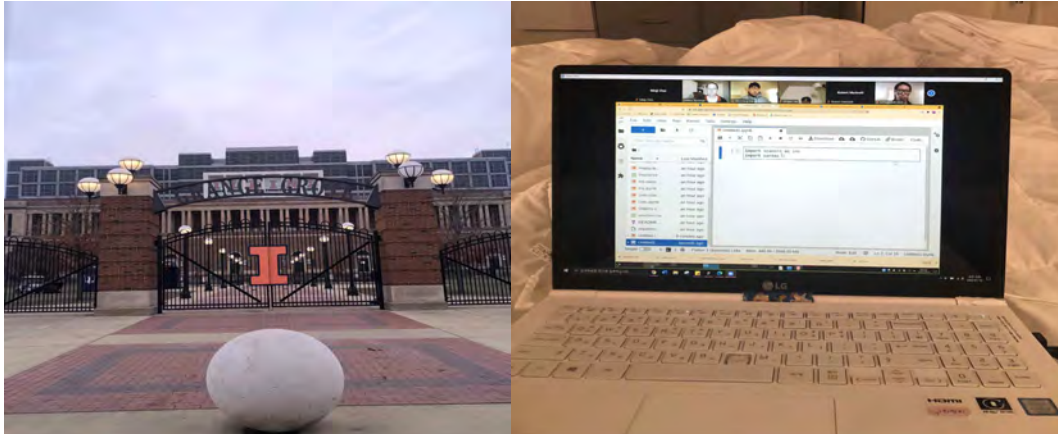
Still however, I thought this program has larger weakness than benefit to me.

Before participating this program I had huge expectations about this program. Because I using the facility and meeting in person with professors in University of illinois which is one of the most famous engineering university was a such rare experience. However, when we become impossible to use and meet school facility, to be honest, I felt really bad and disappointed a lot. I thought doing this was the same as just taking a class with zoom in a hotel in Korea, so the price of the plane ticket was thought to be wasted. Also meeting the presenters with zoom not in person made me feel like just watching the video which eliminated the biggest strength of this program.

- ☐ **Name:** Minji Choi
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

<Lecture>

The UGP-UIUC program was to learn AI and Big data. On the first day of the program, I moved from Chicago to Urbana-Champaign. On the second day, the introduction was conducted, and lectures began in earnest from the third day of the program. Due to COVID-19, all lectures were taken in the hotel room as zoom. Professor Mattox Beckman's lecture, which was the main lecture, learned data visualization by using Python and Jupyter notebook. The most memorable lecture was Peter and Jenny's business English. I was able to learn words, idioms, and small talks used in business situations, and He and she led students to participate through group activities or games, helping them concentrate despite their online lectures. In addition, there were machine learning, data mining, data cleaning, chat bot AI, and introduction to research institutes with professors and ph.D students.



On the 11th day of the program, project presentations were made for each team. Our team analyzed wine quality and pizza data. From wine dataset, we created one main question that what variable have high effect to determine quality of red wine. We could answer the main question, what variable have high effect to determine quality of red wine, as alcohol. Moreover, with the main question of "Which variable affects pizza prices," we analyzed what affects pizza prices through pizza dataset. Surprisingly, meat and sausage did not have a significant effect on the price increase. Finally, we could easily know the price and quality of food predictions through visualization. Some

<Chicago>



On the evening of the 12th and 13th days of the program, a free trip to Chicago was held. I visited Navy Pier and ate Chicago pizza. While walking around Chicago, I could enjoy the night view of building, cloud gate, crown fountain, ice rink, and Chicago theater.



On the morning of the 13th, I visited snow-covered Cloud Gate and Crown Fountains and participated in the Chicago architecture tour. In the afternoon, I visited the Chicago Museum of Art to appreciate the work. For my last schedule in Chicago, I visited a Nandos restaurant that I wanted to eat in the United States. It was not only quarantine due to COVID-19, but also a time of personal regret. But at the same time, it was definitely a meaningful time in that I was able to challenge something and learn new things by myself. It became a process of living alone in an unfamiliar space, looking back on myself objectively, and getting to know myself more. In addition, I was satisfied as I was able to visit the places I wanted to try and go to and carry out the things I planned. Although it was a completely different situation than

expected, I was able to develop my vitality, independence, and even patience through this program, and I am so proud of myself for enduring well in an unfamiliar environment.



- ☐ **Name:** 박수아
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2022-01-03 ~ 2021-02-16

A lot of unexpected things happened in the U.S.

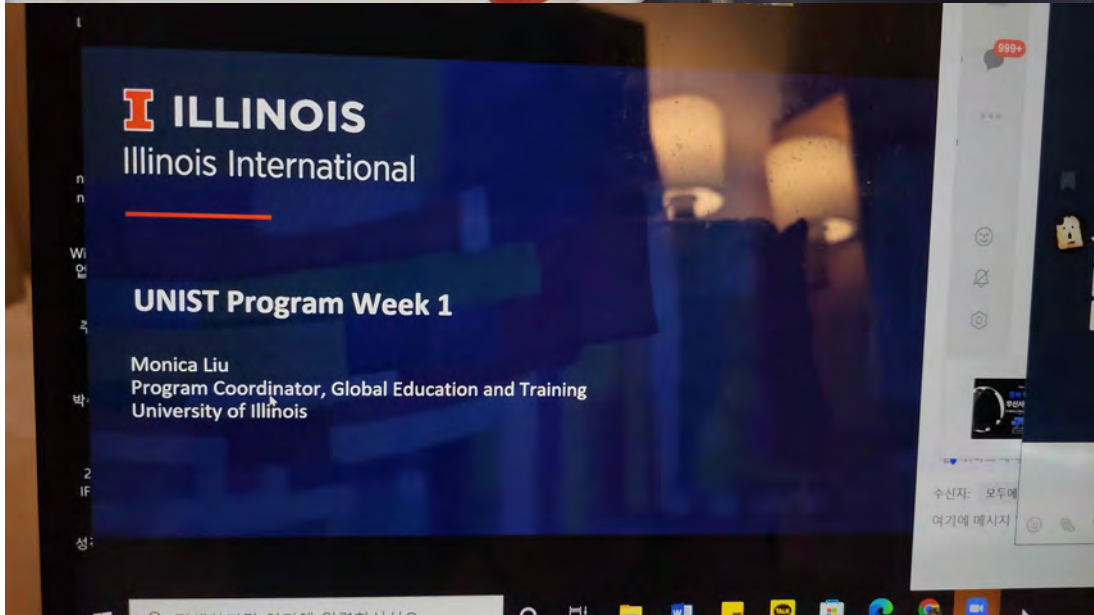
SOO AH PARK

After arriving in Chicago, the camp began with breakfast the next day. First of all, I received the news that I had to self-isolate before the start of the camp, so I received an orientation for the class through the zoom program. I briefly explained about UIUC school and what classes and schedules it will be in the first week. I also tried to adapt by listening to various apps and various help for two weeks of life.

UIUC conducted several PCR tests during this camp through a new PCR test method that has not been experienced in Korea.

At first, I took a class on introducing simple DATA SET and MACHINE LEARNING. I was in a situation where I couldn't take any other classes related to computer science other than the programming class I took in the first grade. Also, I did not major in the department related to the subject of the camp. Therefore, it was difficult to understand 100% in classes that required basic skills. In addition, there were many terms and processes I heard for the first time, so it was difficult to adapt at first.

The DATA VISUALIZATION class was quite interesting. It wasn't a graph that we usually encounter with numbers or easily grasp. This class was a process of showing the results through various graphs so that it was easy to grasp at a glance. At the end, the project was announced. Our team's theme was to VISUALIZATION using data from pizza and wine in India. Various graphs were created with new criteria to find out various things such as topping and price differences according to size. It was more fun to work on the project with topics close to real life, and it was also interesting to see other people's topics.



Also, the second most interesting class was the business English class. I have witnessed a lot of small talks in the United States. I also felt the difference in culture between the United States and Korea for a few days. However, I was not sure if these were American culture. In this class, the professor selected topics that we might be curious about and told us. It deals with basic and confusing things such as greetings from the United States, topics of small talk, and how to call people. It was not a learning part, but a class where I learned more about American culture. I wanted to take more of these classes.





A Global Leader in Higher Education

- Founded in 1867
- 44,000+ students from all 50 states and more than 110+ countries
- 15 colleges & instructional units
- 24 Nobel prizes & 26 Pulitzer prizes
- 13th ranked public university
- Illinois ranked among top 10 Schools for Silicon Valley Recruiters
- One of the largest public university libraries in the world, with 24 million volumes & publications

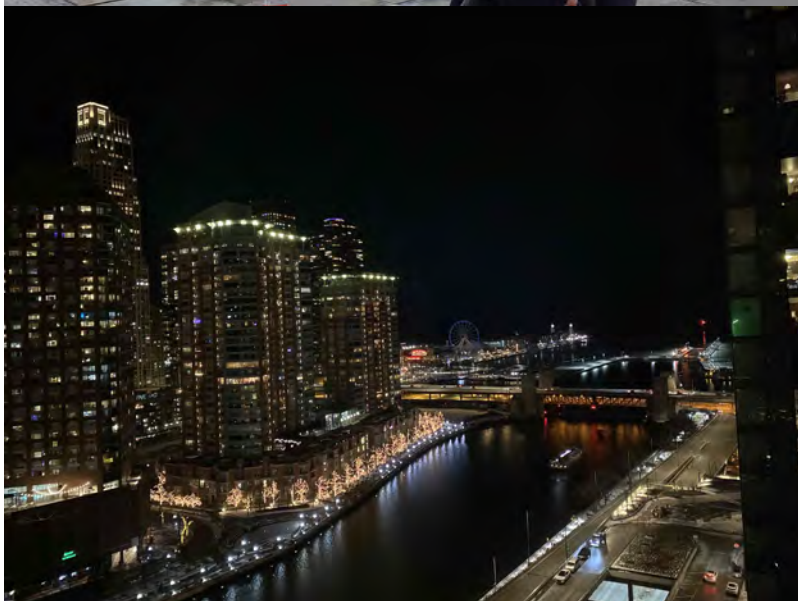


I didn't officially go on a UIUC campus tour, so I went around the campus on my own on the last day. I think about this part very sad and not good. However, since it was my first visit to the U.S., I went around the school with regret.

Also, I visited downtown Chicago to see various buildings. While being tied to a hotel for two days, I visited a famous place in Chicago, took pictures, and looked around. I also tried Chicago pizza and walked along the river. Chicago, like its nickname Windy City, was very very cold and difficult to get around without beanies and gloves.

I hope there will be a longer period of time like this next time.

It was not an expected camp because I only had online classes learning and experience. But I try to be satisfied with meeting good people and experiencing many things.



- ☐ **Name:** Yijin Choi
- ☐ **Host University/Institute:** UNIST
- ☐ **Participating Program:** UNIST – UGP: UIUC
- ☐ **Program Period:** 2021-01-03 ~ 2021-01-16

UIUC Program

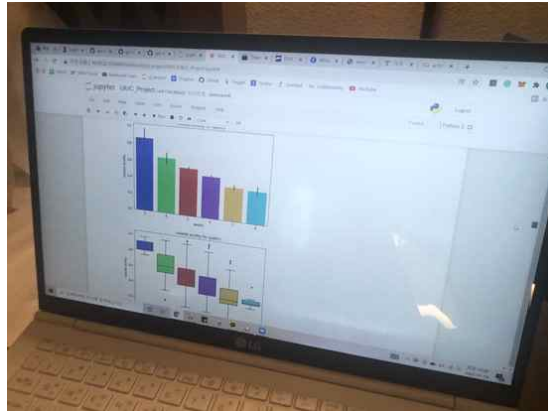
I participated in UIUC program with topic of AI&big data. I will write about program's education, operation, and free travel in Chicago.

1. Education

The program was processed only through zoom because of occurrence of COVID patients and UIUC professors' preferences. However, to say lectures' education here, I will not consider this zoom circumstance and only consider lectures' contents itself.

The lectures are divided into three main categories: AI&big data related lectures, English related lectures, and facility related lectures. And at least for me, I was a little disappointed with them.

First, for AI&big data related lectures, these lectures were why I participated in this program. I expected these educational lectures with thinking that those might be lectures that cannot be learned in UNIST but in UIUC. However, I was a little disappointed with these lectures. The number of these lectures was smaller than I thought and moreover, all lectures were just scratching the surface without the depth. I know there was a time limitation to process professional lectures, but I think the program should be prepared with minimum depth. And I think the program should be prepared with considering students' knowledge level and lectures they learned in UNIST. The main activity was a data group project using python and to do this, mainly one professor prepared lectures related to data visualization using python. However, there were only three lectures related to the project and they were easier than lecture named AI programming that all students in UNIST already took in first grade. I think this made the low quality of group project because students should utilize few plotting methods that they learn in lectures and it was hard to utilize methods that I learned in UNIST for high quality because there might be few students who learned advanced python methods in UNIST. There was some survey related to students knowledge level before the program and I cannot know other students reply to that survey but I think the program should be prepared with a higher level with considering students common curriculum in UNIST. For the next programs, if programs want to plan group projects, I think preparing more lectures related to group projects with appropriate levels and variety of methods might be good to improve the quality of group projects and programs.



And for English related lectures, I thought those three lectures were normal. However, for facility related lectures, I thought there were too many lectures. Without considering study abroad for graduate school in UIUC, honestly, lectures did not come to me well. I think a program with more AI&big data related lectures and fewer facility related lectures might be better for me.

2. Operation

As mentioned above, the program was processed only through zoom and all decision was unilaterally notified to students without some advance notice before the program. As the COVID situation has become more serious right before the program, it seems that the occurrence of a confirmed case could be considered as a possible case during the program and I think countermeasures should be prepared before. However, there was no detailed notification before and after a confirmed diagnosis was made, the related quarantine notice was made 3 to 4 days later. In addition to unilateral quarantine notice, entire zoom lectures were later informed. This raised question about whether countermeasures related to COVID were properly implemented because with considering all possible cases, the occurrence of the confirmed case was possible during the program and if they thought to notice quarantine and zoom lectures to students with considering this case, there was no need to ask students to come to the US. The program could just run zoom lectures from the beginning without making students to pay for expensive program costs and flight tickets. However, it made students to come to the US and with unilateral notification, it made students to stay only in the hotel and have all lectures through zoom that can be done in Korea and it was disappointed. I think the insufficient response to this special situation of COVID has lowered the quality of program operation.

3. Free Travel

After all lectures, there was free travel in Chicago downtown for 2 days.

During this time, I visited the Art Institute of Chicago. The Art Institute of Chicago is one of the popular art museums in the US and worthy of this reputation, I could see many works of popular masters such as Pablo Picasso, Auguste Renoir, Vincent van Gogh, and Oscar-Claude Monet. Through this, I could cultivate artistic knowledge by thinking about the differences between the works of each master and by inferring the master by looking at the other works. It was a good time to have a new experience that I could not have easily in Korea and I could have an interest in art from this time.



And in Chicago where is famous for its architecture, when walking around the downtown, I could look at many architectures of Chicago. I could feel the beauty of the buildings with looking at architectures such as the Bean and Marina City Tower Building that can only be seen in Chicago.



- ☐ **Name: Eunseok Park**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**

The effect of COVID-19 in education across the world.

Eversince the admission into UNIST, one of the things I had in mind was to study abroad for graduate school or post doctoral scholar. Luckily, I received an opportunity to experience the education of the United States, as I got a scholarship for the UGP, or the UNIST Global Program. I was given two choices to choose from: Aalto University at Finland or UIUC at the United States. From these two, I chose UIUC to study AI and Data at the school of informatics, as United States was a place I wanted to study in. The reason for this can be explained with an experience as a child. I was an introvert when I was young. I couldn't stand up in front of others and was always anxious of meeting new people. But this changed totally when I got a chance to live in the United States. Living several years in New York, I experienced the culture and education of the US. Eversince, my personality was a complete turnaround. I became a very outgoing person and was willing to talk to new people on the street. This experience was a very inspiring one in my life and have been desiring to go back to the United States henceforth.

I planned on going on a personal trip prior to the schedule once I was given a chance to go to UIUC. The city of Chicago, also known as the windy city, is known for its Willis Tower, the highest building in the midwest United States, the Cloud Gate, as known as 'The Bean' and the Museum Campus near the Lake Michigan, one of the Great Lakes. After getting out of the Chicago O'Hare airport, it was a challenge to find the way to the city. Unlike Korea, with a well organized metro system near the airport, I had to find a a shuttle to the subway station. But the first encounter of Chicago was not very expected. The moment I got out of the metro station of Chicago, I was able to



The Willis Tower

smell a nose-stinging odor which turned out to be a marijuana. I was surprised, as I thought it was supposed to be illegal in the state of Illinois. But otherwise, the buildings and the weather of Chicago was beyond of what I thought of. The harmony of buildings of the 19th century and the buildings of the 21st century was an amusing scene to watch.



The City View of Chicago-1



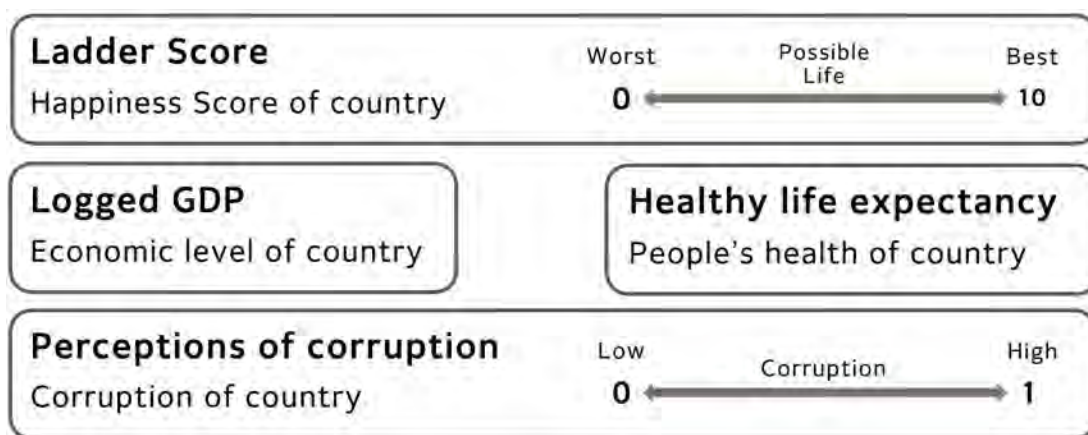
The City View of Chicago-2

But the nickname of the city was not a joke for sure. The icy wind blowing from the Great Lakes combined with the snow, formed a blizzard, making it hard to walk through the streets. Through the blizzard, I was able to go all across the city to visit the landmarks of Chicago. After a short encounter with Chicago, I went back to the airport to ride a bus with my fellow students from UNIST to Champaign, with a schedule at the UIUC, which I was looking forward to.

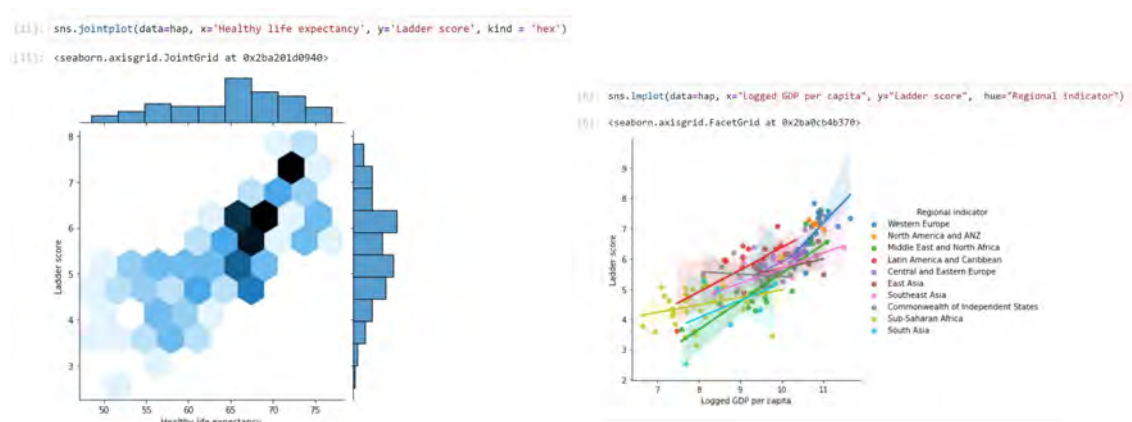
However, the problem broke out then. Some of the students that also went on a trip before the program were tested positive on the COVID-19 test the day after we arrived. Because of this, all the students that went on the same bus were regarded as a person with close contact to the infected. We were forced to quarantine in a hotel room, and not at the actual school class. In a consequence, we had to continue our class via zoom like the ones we used to take at Korea. But the zoom class still gave us sessions about coding that might be useful for the projects we can do in the future, and business english sessions that we can use on interactions with the foreign researchers in the future.

From the short explanation of coding tools Seaborn and Jupyter, we formed ourselves into groups to work on a small project regarding a network of lots of

things over the world. From a lot of networks, we were interested in the people's life satisfaction. We were able to get well-being and lifestyle data, and world happiness report in 2021 from Kaggle. In a categorical data, there are country name and regional indicator such as western Europe, east Asia, and the numerical data that go with them. The following figure shows the representative numerical data. The ladder score is the happiness score in a 0 to 10 scale, 10 as the best and 0 as worst. Logged GDP per capita indicates the economic level of the country. Healthy life expectancy indicates the people's health in each country. Perceptions of corruption indicates the corruption of each country. In this dataset, as the perception of corruption of country is close to 1, the corruption of country is high and as the perceptions of corruption of country is close to 0, the corruption of country is low.



We used the scatter type to plot some graphs. Additionally, we tried to make the jointplot using hexagonal type. The result is this plot. Comparing the jointplot in the previous slide, hexagons exist and the color of them is different. The dark hexagons mean more distributed variables. We think the hexagonal type is more effective to express the distribution of variables. Expressing by color difference is more intuitive. The plot following is using lmplo. In lmplo, we can find the relation between two variables using linear regression. And also, we can introduce hue. So, we can express categorical data in different colors in the plot.



In this plot, we tried to find the relation between ladder score and logged gdp per capita depending on region. In this plot, we can find all linear regression plot is increasing. It means in all regions, higher logged gdp per capita makes higher ladder score. Also, we can verify again that the western Europe region has the highest ladder score and the slope of the linear regression plot is the highest. From these graphs, using the world happiness report 2021 dataset, we visualize the data and analyze the relation between variables using seaborn.

Through these projects we were able to learn a bit about informatics, using seaborn and jupyter, and creating network analysis with the datas from Kaggle. From these short lessons, we were able to learn a bit about American culture and the classes of American universities. I think it has been quite an experience in terms of learning the culture, despite the zoom classes and the limited time.

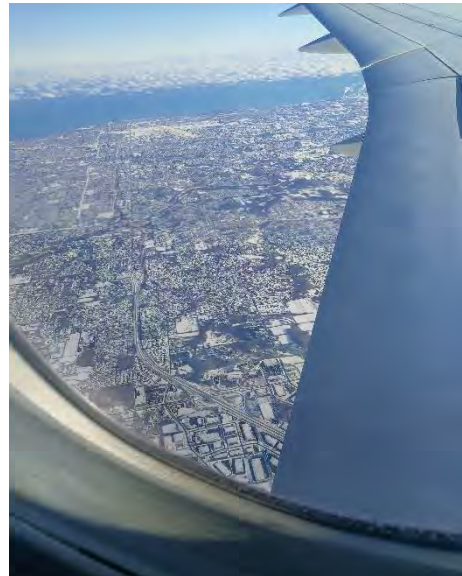
- ☐ **Name: Taeyeong Bak**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**

This was my second trip to the US and my first study abroad program. On the first day I decided to go to America, I didn't realize that I was going to America. The first time I went to the United States was when I was in high school and had to plan and prepare everything by myself. Everything I did on my own, such as finding plane tickets to the United States or deciding where to go in downtown Chicago, was a new challenge for me and made me grow.

A lot has happened since the journey to America. I did not take a direct flight to the United States, but went via the United Arab Emirates. The flight was delayed due to the airline's error, but I did not receive a notice of the delay and waited for 2 hours. also responded. Even after arriving in the United States, the plane arrived late, of course, and I waited for all my baggage for two and a half hours, but I did not see if my baggage had been lost. Fortunately, I was able to retrieve my luggage the next day, however, having gone through all these things, I learned that traveling alone exposes me to various risks at any time, and I must always be prepared because these things may happen to me as well. Also, whenever

I go through a crisis like this, I have to deal with it, and, to a lesser extent, my English skills have improved as well.

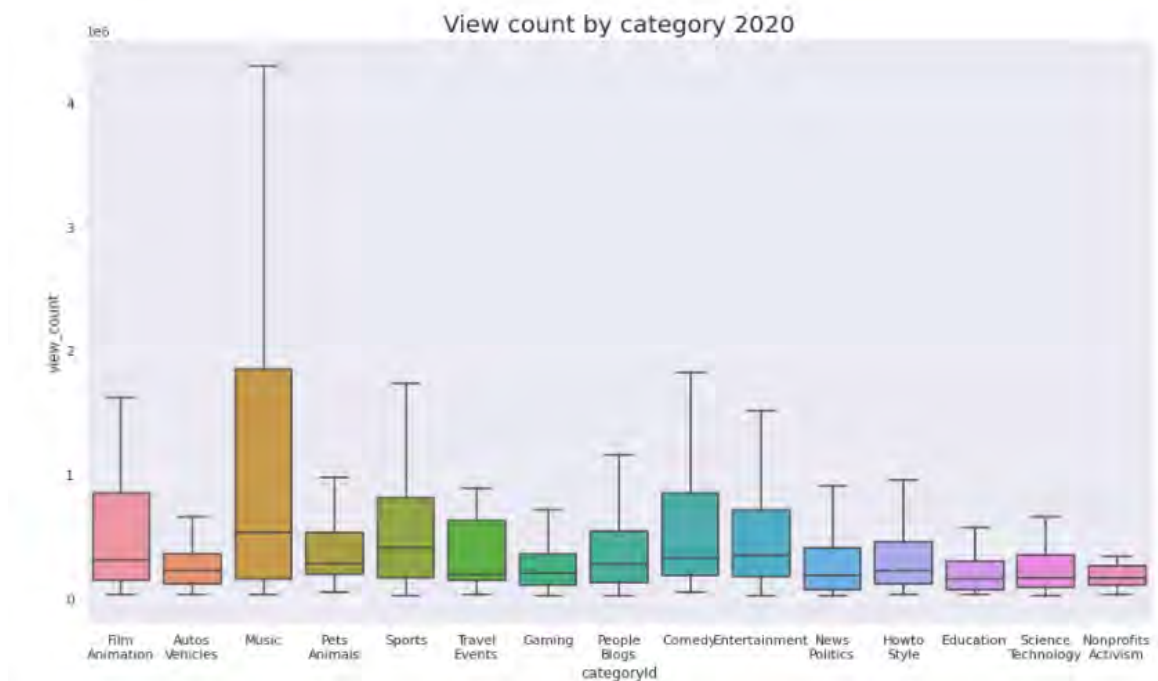
Even so, I couldn't help but lose my heart from the appearance of America I met on the plane. The scenery of the snow-covered foreign city from above the sky was so beautiful, it was as if I was being compensated for the hardships I had been through.



On January 3rd, when I checked in at the hotel, I thought that it was finally the beginning. We met UIUC professors online and were able to attend several lectures on data science and artificial intelligence. All of the lectures were excellent, and they helped me a lot. At the meeting with UIUC graduate students, I was able to hear vivid experiences about what university life in the United States was like, why I chose the difficult path of studying abroad, and what preparations were made to study abroad. All the lectures, including the Korean class, were memorable and very helpful.

But, really sadly, as one of my co-workers tested positive for coronavirus, the class ended up being held only online until the last day. I'm in my 20s, so I've never had a face-to-face class yet, and I know that the fact that classes with UIUC

were only conducted online was an unavoidable situation and appropriate response, but it was really disappointing. Also, I used the hotel room as a single room, and while it was really comfortable and enjoyable for two weeks, I didn't have many opportunities to get to know people.

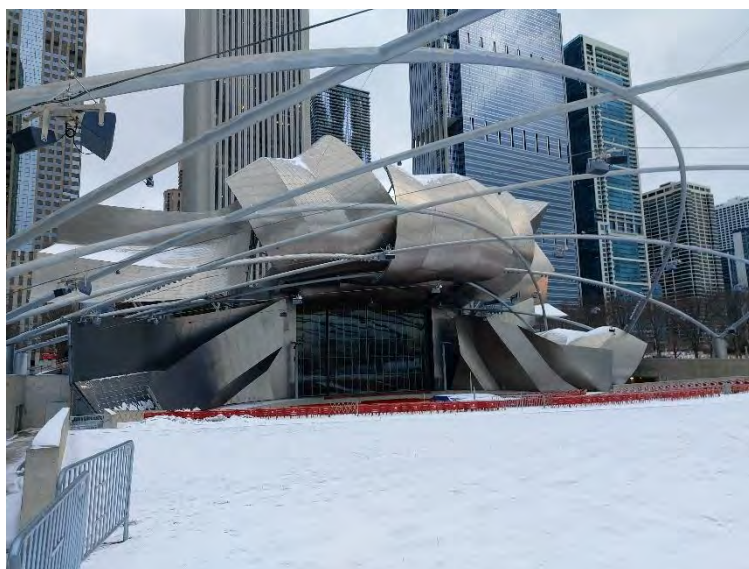


But even in this difficult situation, the final project connected us. While preparing for the final project, I got to know other people, and the process of preparing for it was really fun. The most memorable part of this schedule is to think about what you have learned so far, find interesting data, and process and interpret it in your own way. The above photo is a data I made by writing the code myself, and I analyzed the daily popular video data on YouTube and presented it by adding various analysis criteria such as year, category, like/dislike ratio, etc. and my own interpretation.

It may be due to the self-quarantine, but those two weeks passed much faster than I thought. Although it was very disappointing that it was online until the end, I have no regrets about the people I met through this program and the things I learned.



After class, for the last two days remaining in the United States, I visited downtown Chicago and went on a free tour. In fact, until I came to the United States, I had no idea what Chicago was famous for



or what to see. Even in Korea, when I go on a trip, it is not the style of planning and preparing, but even more so. However, it was really fun to find and schedule tourist

attractions on my own, given the advice of thinking about what to see and experience. In such a short time of two days, although I did not experience 100% of what I had planned, I did my best to play as hard as possible and experience many things, and these two days are likely to remain in my memory the most.

I would like to express my infinite gratitude to UNIST and UIUC for planning such a good program, and I hope that those who are going to study abroad in the future will have better experiences and more experiences than us.

- ☐ **Name: Dongseok Yu**
☐ **Host University/Institute: UNIST**
☐ **Participating Program: UNIST – UGP: UIUC**
☐ **Program Period: 2021-01-03 ~ 2021-01-16**

UIUC UGP Program Final Report

Introduction

In this report, I would like to write about what I have learned and felt through participating in the UIUC UGP program. The official topic of the UGP program was AI programming, but the actual program included other interesting topics such as business English and the 4th industrial revolution.

Data analysis with SHAP mechanism

As stated, I used a machine learning explainability method to understand a machine learning model's behavior. The machine learning model predicts a learner's final test score (posttest score) based on the various features of the learner's behavior.

There are three types of features:

- Pretest score
- Quiz score (0-100)
- Reading time (ms)

Each group is assigned one dataset. I analyzed data with various features. For example, how many features are there for each type? How many observations are there? Are there any missing values? What might missing values (or -1) mean? What could be the context for the data?

The most important thing that I learned was utilize shap to visualize and interpret the model based on the data. I was able to see how features were influencing the predicted posttest score, as well as the direction of the influence (whether it negatively impacted or positively impacted the posttest score).

First, I looked at the numbers for each feature to see the organization of the data.

	participant_id	pre_test_grade	post_test_grade	quiz_grade_1	quiz_grade_2	quiz_grade_3	quiz_grade_4	quiz_grade_5	quiz_grade_6	quiz_grade_7	quiz_grade_8	quiz_grade_9	quiz_grade_10	quiz_grade_11	quiz_grade_12
0	learner16	41.666667	50.000000	33.333333	100	100.000000	66.666667	66.666667	0.000000	33.333333	66.666667	66.666667	33.333333	66.666667	33.333333
1	learner17	25.000000	66.666667	83.333333	100	66.666667	100.000000	66.666667	100.000000	33.333333	100.000000	66.666667	100.000000	66.666667	100.000000
2	learner18	41.666667	75.000000	66.666667	100	100.000000	66.666667	100.000000	33.333333	66.666667	66.666667	66.666667	100.000000	66.666667	100.000000
3	learner19	66.666667	91.666667	33.333333	100	100.000000	66.666667	100.000000	100.000000	66.666667	66.666667	-1.000000	-1.000000	-1.000000	-1.000000
4	learner20	50.000000	75.000000	33.333333	100	66.666667	66.666667	33.333333	33.333333	66.666667	100.000000	100.000000	0.000000	0.000000	0.000000

In this section, I found that this data quality is pretty bad because there are some non resonable values such as -1, or 0.

And then I used follow code to see SHAP values.

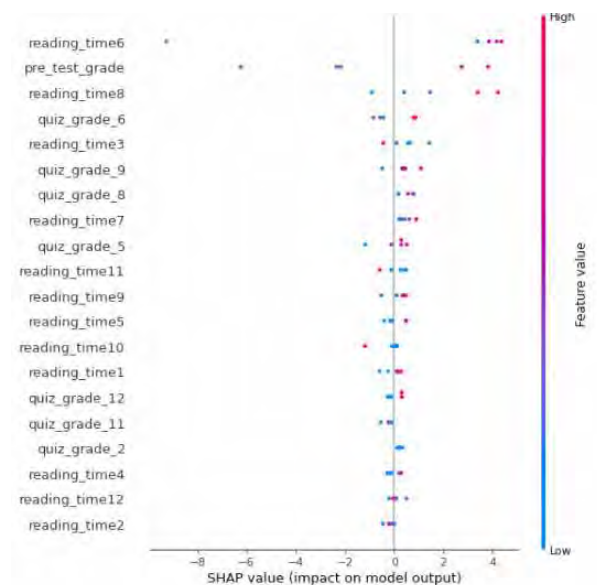
```
feature_names = ['pre_test_grade',
                 *['quiz_grade_' + str(v) for v in range(1, 13)],
                 *['reading_time' + str(v) for v in range(1, 13)]]

learner_features = df[feature_names]

explainer = shap.TreeExplainer(model)

shap_values = explainer(learner_features)

shap.summary_plot(shap_values)
```



And then I made a base value and found the negative and positive effect on their learning with follow code.

```

-----
shap_values = explainer.shap_values(learner_features)

list_shap_val = shap_values[(for trainer)]

list_shap_val

shap_df = pd.DataFrame({
    'feature_name': feature_names,
    'shap_value': list_shap_val,
    'feature_value': learner_features.values[4],
})

shap_df

shap_df = shap_df.sort_values(by=['shap_value'])

shap_df

shap.initjs()

shap.force_plot(explainer.expected_value, list_shap_val, feature_names = feature_names)
-----

```

And I compared a positive and negative effect on the other trainer.

With this activity, I learned how to analysis a complex data with machine learning. And I used this skills for the final presentation.

Final Presentation

We are interested in the people's life satisfaction, which is recognized important in today. First, we chose well-being and lifestyle data in Kaggle. However, this data is composed of the survey. And all variables are all 1 to 5. Thus relation analysis of data was hard. So, we chose another dataset of the recent world happiness report in 2021 data in Kaggle.

In categorical data, there are country name and regional indicator like western Europe, east asia, north America and so on. Also, there are lots of numerical data. Ladder score is the happiness score in the world as 0 to 10 scale. The 10 is best possible life and the 0 is the worst possible life. Logged GDP per capita indicates the economic level of the country.

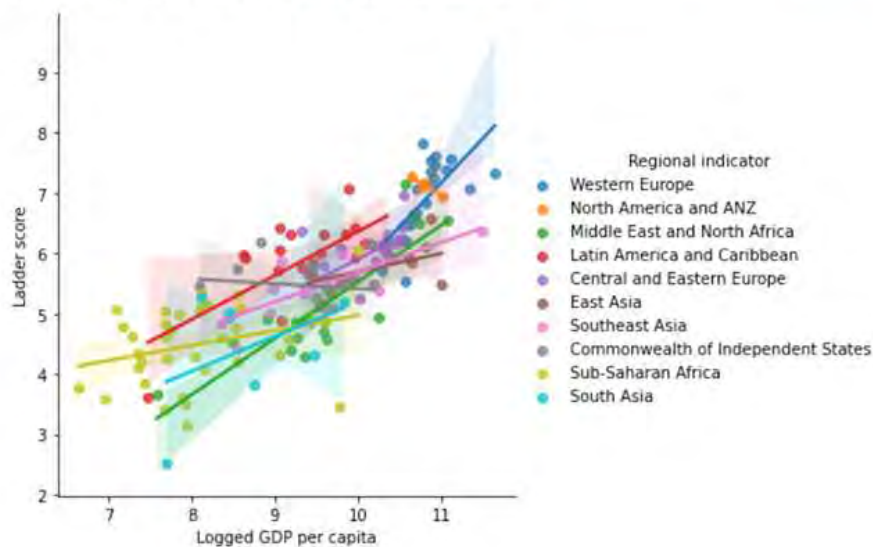
Healthy life expectancy indicates the people's health in each country. Perceptions of corruption indicates the corruption of each country. In this dataset, as the perception of corruption of country is close to 1, the corruption of country is high and as the perceptions of corruption of country is close to 0, the corruption of country is low.

In lmpplot, we can find the relation between two variables using linear regression. And also, we can use hue. So, we can express categorical data in different colors in the plot. In this plot, we tried to find the relation between ladder score and logged gdp per capita depending on region.

And also, we can find all linear regression plot is increasing. It means in all regions, higher logged gdp per capita makes higher ladder score. Also, we can verify again that the western Europe region has the highest ladder score and the slope of the linear regression plot is the highest.

```
[6]: sns.lmpplot(data=hap, x="Logged GDP per capita", y="Ladder score", hue="Regional indicator")
```

```
[6]: <seaborn.axisgrid.FacetGrid at 0x2ba0cb4b370>
```



The plot that we found is upper figure.

Finalizing the report

We are interested in the people's life satisfaction, which is recognized important in today. First, we chose well-being and lifestyle data in Kaggle. However, this data is composed of the survey. And all variables are all 1 to 5. Thus relation analysis of data was hard. So, we chose another dataset of the recent world happiness report in 2021 data in Kaggle.

UIUC Report

- ☐ **Name: Kwangrok Kim**
- ☐ **Host University/Institute: UNIST**
- ☐ **Participating Program: UNIST – UGP: UIUC**
- ☐ **Program Period: 2021-01-03 ~ 2021-01-16**

형식

- 본 파일 / 최소 3장
- 제목/내용/사진등을 포함한 소감을 적어 자유롭게 작성

UNIST 공식 인정 단계

- 파견 완료 공식 인정 절차
- 보고서 미제출 시, 해외연수장학금
(프로그램 비용, 숙박비, 장학금 차액 UNIST로 반환)

제출 기한

- 프로그램 완료 후, 2주일 이내로 전원 작성 후, PDF 스캔본으로 제출
- 2021.02.13 (일)까지
- 개인 여행 여부와 관계없이 제출

I participated in the UIUC AI course for 2 weeks. The program was conducted as a cyber lecture with self-isolation because some students had COVID-19. So it's a pity that I couldn't meet the professor in person and use the school facilities. However, I think it was a good time because I could learn about ai. After self-isolation, I was able to go out for about two days. So I looked around the school. I went to facilities here and there with my friend and took a lot of nice pictures. Also, at the end of the program, I had time to spend freely in downtown Chicago. I went to the museum with my roommate and went to a jazz bar while looking at the night view.











