Brandon Roten

INFO 3401

Problem Set 3

**Monday:**

**1. Use the command line to run asciify.py (included in the assignment repo). What does this script do? Include a screenshot of your results.**

This script prints characters that create an image of a cat. Screenshot is below:



**2. Install the Delorean Python package. What command did you use? Then, from the command line, launch Python. Use it to execute the following commands:**

from delorean import Delorean

EST = "US/Eastern"

d = Delorean(timezone=EST)  
print(d)

**What prints to the command line?**

Delorean(datetime=datetime.datetime(2018, 9, 16, 21, 7, 30, 605448), timezone='US/Eastern')

**3. Use the command line to download**[**http://www.colorado.edu/**](http://www.colorado.edu/)**. What command did you use? Make sure to include this file as part of your submission.**

I used the “wget” command.

**4. Search the downloaded file from #3 for the word "Colorado." What command did you use? How many lines did this find?**

I used the command “grep.”

**5. Now, print the lines containing the word "Colorado" to a file to "search\_results.html". What command did you use? Include this file as part of your submission.**

I used the “cat” command.

**Wednesday:**

CU has recently come across a variety of findings stating that regular gym attendance can lead to a higher GPA. In an effort to understand this effect and how they might update their services and facilities to encourage students to be active, they want to turn to the data to both understand the current state of gym use, its potential effects on students, and how they might target specific facilities or programs to improve attendance.

First, identify three **hypothetical**datasets that CU can provide you to support this effort (you have a sense of what kinds of data they may have available, just describe what data you would use). Then, use the CoNVO method to scope a project that will solve this problem. Provide sufficient detail in your response such that someone could reason about and execute your plan.

Data Set 1: Gym vs. Study Comparison

Co – People who would be interested in this information would be people who go to the gym, as well as people who don’t. This study is trying to show that people who take breaks from studying to go to the gym end up with a positive result. These results would make students pause and consider going to the gym instead of just studying.

N – This will help people want to become more active to benefit their schoolwork, with a bonus result of becoming more healthy. Those who already attend the gym will be pleased to see the results and continue their routine.

V – To achieve this, the data would need to be collected on all students, being those who study and attend the gym, and those who just study. There will also need to be additional data used involving success in school.

O – These data results will be used to encourage students to combine both studies and the gym in order to receive better results in school. The next step after using this data might be to pass it along to other universities and compare/contrast results.

Data Set 2: GPA of Gym vs Non-Gym

Co – People who would be interested in this information would be both students who attend the gym, and those who don’t. People who promote the gym would also want this information, because they could better motivate students to go to the gym.

N – The results will make students want to go to the gym more because they want to improve their grades. Sometimes students may think that they would benefit from doing other things rather than the gym for the GPA, so this would be impactful information.

V – This data research can be achieved by gathering GPA results, as well as running a study on which students also go to the gym. Then this would be cross-examined, and then used to display the results.

O – The results of this study would lead students to really consider going to the gym in order to boost their GPA. The school itself would also be able to promote a more healthy lifestyle for students because it also leads to more success in the classroom.

Data Set 3: Active Brain Study

Co – There would be many people interested in this study, because it could be potentially lead to a change in how students all around the world operate. This study could be eye-opening to how the relationship of the body and mind work.

N – The results of this study will change the way that students operate on a daily basis. As mentioned previously, it can change how we understand the mind and body.

V – This study would be performed by gathering subjects, some that go to the gym and some who do not. There will then be data collected on if having an “active body” can lead to having an “active mind” as well.

O – The results of this study would easily help any institution promote students going to the gym. This data would be passed along to various universities and organizations, leading to deeper study.

**Friday:**

In class today, you learned about a historical instance of data science in practice. Find three additional problems in history that were addressed using data. For each of these problems, describe the following:

*1. What was the problem?*

UPS used to replace parts on all vehicles on a routine basis, which was very expensive.  
*2. What data was available?*

The data that was available at the time was the cost and parts used.  
*3. What methods were used?*

They placed sensors on all of their vehicles that allow them to know how the parts are performing. They are then put through an algorithm that determines the performance of the part.  
*4. What did they find?*

They found that they were using too many parts and loosing too much money for the same performance of their vehicles. These sensors are still in use today.  
*5. If you were to tackle the same problem today, what would you do differently and why?*

If I were to do this again today, there isn’t much I would change. If anything, I would have just done this sooner.

*1. What was the problem?*

Humans were trying to reach the moon for the first time. *2. What data was available?*

The data that was available was from previous attempts and simulations of flying the space shuttle.*3. What methods were used?*

Specifically the Apollo Guidance Computer was coded with MAC which was then converted by hand to assembler language that the computer could understand. The system was meant for the astronauts to be able to use simple commands to operate the shuttle effectively. *4. What did they find?*

The mission was a success, but this creation was more basic than a modern toaster. It was said to have 64Kbyte memory, and operate at .043MHz. *5. If you were to tackle the same problem today, what would you do differently and why?*

If I was to tackle the same problem today, I would obviously use much more advanced code that would allow for swift connection, and easy operation.

*1. What was the problem?*

The U.S. had a concern for the safety of the Country. *2. What data was available?*

Pretty much just about any data that existed on the internet. *3. What methods were used?*

Large scale data scraping, and data mining. *4. What did they find?*

Unclear, but purpose was to discover acts of terrorism. *5. If you were to tackle the same problem today, what would you do differently and why?*

If I were to tackle this problem today, the main thing I would do is be more ethical, and strict with the data scraping.