

9-16-2018

Monday

1. It converts images into ASCII.

[illegible]

2. I used "pip install Delorean"

```
Type "help", "copyright", "credits" or "license" for more information.
>>> from delorean import Delorean
>>> EST = "US/Eastern"
>>> d = Delorean(timezone=EST)
>>> print(d)
Delorean(datetime=datetime.datetime(2018, 9, 16, 18, 46, 58, 930774), timezone='US/Eastern')
>>>
```

3. n/a
4. I used "grep -n Colorado index.html" First, I used "grep Colorado index.html" but it made it hard to count the number of lines. It found 17 lines.
5. I used: "cat *.html | grep Colorado > search_results.html" I only have one .html file on my machine or I would have not been able to use "cat *.html"

Wednesday:

Hypothetical Datasets

By using the data of a student's ID card, we could come up with the following hypothetical datasets:

1. Those students who go to the gym often, as in 4x a week or more
2. Those students who never go to the gym
3. Students who are casual gym users, 1-2x per week

Context

- CU Students are the stakeholders
- We are trying to use data to see if there is a correlation between high fitness center use and a high GPA
- This project will advance CU's knowledge of what other outside forces have an impact on GPA's
- The target audience is any current/future student in any field/college

Needs

- One problem that could be fixed with this data is getting student more accustomed to using the fitness center at CU if they are not familiar. Even if we find out that fitness doesn't affect GPA's, personal fitness is still a worthwhile cause.
- We hope to know if frequent fitness center use causes a higher GPA amongst students

Vision

- Some techniques we could use are interviews, surveys and observations, student ID card data
- By using interviews in conjunction with ID card data, we could go deeper than data at surface level to find out if GPA are affected by fitness
- When we are done we should have a better character profile on students who go to the gym often/casual gym use/no gym use and the correlation on GPA's.

Outcomes

- Our results will be used to better guide future/current students as to how they can better perform
- It will be integrated back into the organization as a way to better educate future/current students to get a more rounded education and to help them better understand how fitness can improve other aspects of their life.
- When we are done, if we have proven that fitness does affect a student's GPA, we can start building a more targeted approach to get students to be more fit.

- We could pass this information at to a 3rd party who has a better understanding of how to get students to better utilize the fitness options offered by colleges.

Friday:

Problem #1:

1.What was the problem? Improving public safety. In Memphis, they wanted more police on the street, but they were often stuck doing “grunt work and shuffling papers.”

2. What data was available? Crime data stockpiled from years of reporting in the Memphis Police Department.

3. What methods were used? Statistics and mapping.

4. What did they find? “The team learned that putting officers in locations and times that matched those offense patterns resulted in either deterring criminals or catching them.”

5. If you were to tackle the same problem today, what would you do differently and why? This system they used was ultimately named the Blue CRUSH system. It started in 1997, long before almost every person carried a cell phone with them. I think it would be interested to go back and re-think what could be done with GPS coordinates for crimes in the Memphis area. I think by using cell phone data, you would be able to get more accurate sense of when and where crime is occurring.

1.What was the problem? Blood disease detection

2. What data was available? Huge medical data collected from many medical tests which were conducted in many domains.

3. What methods were used? Data Mining (Decision Tree, Rule Induction, and Naïve Bayes)

4. What did they find? The results show that Naïve Bayes classifier has the ability to predict the Tumor of blood disease better than the other two classifiers with accuracy of 56%.

5. If you were to tackle the same problem today, what would you do differently and why? I would attempt to expand on other data that was recorded such as climate, location and age to attempt to draw more data points that might point to certain blood diseases.

1. What was the problem? Pruning and maintaining trees in NYC – it is considered a public safety issue. Four people were killed from 2009-2010 in Central Park alone from falling tree limbs – that is crazy!!

2. What data was available? The data NYC had used from years of pruning and trimming trees around the city.

3. What methods were used? A third-party used Data Modeling to build a risk profile. They cleaned, scrubbed and merged all data sets NYC had on their trees’ maintenance schedules.

4. What did they find? pruning trees for certain types of hazards caused a 22 percent reduction in the number of times the department had to send a crew for emergency cleanups.

5. If you were to tackle the same problem today, what would you do differently and why? I would use the data points where fatalities occurred to see if there was some other factors as to why a tree branch fell. e.g. was the tree located in a part of the park where wind was more likely to pass through? Was the tree older relatively speaking to the trees around it?

.