New Group Project Ideas Sheet

Graham, Danieli, Enzo, Katlin

1st meeting: 3/26/22
—--

Recipes

Potential APIs
- https://rapidapi.com/spoonacular/api/recipe-food-nutrition/
- https://rapidapi.com/apidojo/api/tasty/
- https://fdc.nal.usda.gov/api-guide.html
- https://www.programmableweb.com/news/10-most-popular-food-apis-2021/brief/2021/05/05

Interesting Reads…Be Beware…A lot like our idea.
- https://towardsdatascience.com/using-machine-learning-to-generate-recipes-that-actually-works-b2331c85ab72
- https://www.youtube.com/watch?v=CjCNHcb-Br8

—--

Natural Disasters

Read me add-ons:
Types of plastics:
1-PET: **The most widely recycled plastic** in the world is PET. Nonetheless, despite being a relatively easy plastic to recycle.
2-HDPE: HDPE is accepted at most recycling centers in the world, as it is one of the easiest plastic polymers to recycle.
3-PVC: Unfortunately, plastic number 3, the so-called polyvinyl chloride is not recyclable in normal collections.
4-LDPE: Low-density polyethylene, also known as plastic-type #4, is used to make the *infamous* **plastic bags like those provided by grocery stores and other retailers. Technically, LDPE can be recycled. However, as previously mentioned, just because something** *can* **be recycled doesn't mean it** *will* **be recycled.**
5-PP: The last plastic-type that can be recycled is PP also called polypropylene. While PP is easily among the most popular plastic packaging materials in the world, only around 1-3% is recycled in the US, which means most PP is headed for the landfill. Here it degrades slowly and takes around 20-30 years to completely decompose.
6-PS (POLYSTYRENE) and 7-OTHER (Finally, plastics number 6 and 7 are never recycled)

<u>**Project Wildfire:**</u>
After reviewing the datasets of previous project ideas the Wildfire dataset was chosen as the project basis. Now with a working dataset we decided to move forward with the wildfire project.

Charts:
- **Line chart of time vs fire size with causes, Manmade and natural. y= time, x= fire size, and the lines would be the fire causes.**

- **Histogram: causes of fire**

- **Scatterplot of fire size and location?**

Notes from meeting on April 23, 2022
- **Database tables can show relations of fire size over time and manmade vs nature causes.**
- **state/month group could result in a target value for linear regression**
- **There is a possible API pull for region mapping of fire zones from: https://data-nifc.opendata.arcgis.com/datasets/nifc::national-ia-frequency-zones-federal/api**
- **Possibly show real estate value in boundaries around fire zones**
- **Fire Size = Acres**
- **Fire size/class possible groups/buckets for unsupervised machine learning.**
- **If we have time try running two different types of models to compare results and draw new conclusions.**
- **The flow chart in gitmind should be used for the ERD and update it to reflect the new cleaned data with new columns**
- **New dataframes were created and pushed to github as CSV with filtered/cleaned data**
- **Database table csv were created**
- **Establish bins, binning code created, and bins checked**
- **New grouping of causes within columns identified (example: Utilities- railroad, powerlines, equipment use.)**
- **Accidents: Children, campfire, smoking, fireworks**
- **Possible additional column- criminal, non-crime (true/false value)**
-

Notes from April 24, 2022

- New dataframe tables created with grouping of causes, Utilities (railroad, powerlines, equipment use), Accidental (children, campfire, smoking, fireworks), Other (missing/undefined, miscellaneous, structure)

- Tested charts for various columns to look for a helpful chart, scatter plot, bar charts
- Conclusions to analyze:
  - Has the number of wildfires changed over time?
  - Has the size of wildfires changed over time?
  - Has the cause of wildfires changed over time?
  - Has the timing (month) of wildfires changed over time?
  - How does vegetation type impact the probability of ignition?
  - How do temperature, wind, humidity, precipitation impact the probability of ignition?
  - Does "remoteness" have an impact on wildfire ignition?
- Perform unsupervised learning model on df30
- Filter the DF30, df 15, and df 7 tables for data analysis comparison

Notes from April 30, 2022
- Columns updated and cleaned to reflect more climate based influences
- Discussed which charts would be beneficial for comparing time versus fire size
- Discussed how databases should

Notes from May 2nd, 2022
- Discussion of potential roles and how to frame current works and examples to fit the framework of the rubric.
- Review latest edition of cleaned data and df creations.
- Database creation and layout:
  - fire_name --> 'fire_index', 'fire_name'
  - fire_info --> 'fire_index', 'fire_size', 'fire_cause', 'discovery_month', 'putout_time','year'
  - fire_location --> 'fire_index', 'latitude', 'longitude', 'state',
  - weather_data --> 'fire_index','Temp_pre_30', 'Wind_pre_30', 'Hum_pre_30','Temp_pre_15', 'Wind_pre_15', 'Hum_pre_15',
  - 'Temp_pre_7', 'Wind_pre_7', 'Hum_pre_7',
  - fire_city --> API--> "city", "fire_latitude", "fire_longitude"
- Review Unsupervised Machine learning clustering models
- Links for website: Github. Tableau, and google slide show presentation
- Viewed website layout

Notes from May 3rd, 2022
- Reviewed work thus far and submitted to instructor
- Plan on binning to use supervised learning models and add prediction of fire size
- Stats to add to website: largest fire, most common locations, and common fire conditions
- Queries for database: "What are the top 5 fire sizes?"
  - "What state has the most fires?"
  - "What year has the most fires?"

- Possible relation of weather_data to highest fire size?

Add on for 3rd May

- Start thinking about grouping by conditions (temp, wind, hum)
- Weather data by location to predict fire in specific locations.
  - Or predict the size of the fire
  - Target would be fire size for supervised ML. (Does not need to be 1&0)
  - Option to use- OneHotEncoding

- Cannot use Logistic Regression
  - Use Linear Regression and/or Neural Network

Notes from May 10th-11th, 2022
- Try to fix the issue with the SQL/python errors in creating new tables.
- New tables cannot be created using joins and INTO
- Database is searchable for queries
- "The reason predicting whether or not a fire will grow beyond a "Medium" size is to help assist with asset relocation, early containment, and any possible evacuation. This will also help first responders and is a useful tool for local fire services to gauge whether or not they need to closely monitor certain fire activity."
- Determining fire size is essential to current trends as the climate change crisis now creates conditions that are only deteriorating and causing worse destruction. With this fire prediction tool we hope to provide new information that can help assist fire prevention teams with distribution of assets and even help them to request more resources from local and federal government programs.
- Using the tableau map visualizations will also help determine areas that are highly affected and need attention on a growing basis.
- Climate change affects us all and needs to be a topic of discussion for levels of government. As voters and the people who are affected the most the general public needs tools like this that bring the issue to the forefront and make it easy to understand so we can all come together and work together in the fight for a more stable future.
- (bin_count.png) image shows the number of fires and the reclassification of fire size to help illustrate the importance of containment and how frequent small fires are. This also goes to show that with proper education the number of non-natural fires can be reduced.
- Ideally together we can help to stop the destruction of a wildfire and prevent future loss of life and reduce property damage.

Notes from May 12, 2022

- Title and subtitle for website, drop introduction paragraph a line or two to provide more space for image slideshow
- Background sub heading from readme included underneath the intro paragraph
- Scorched slideshow (left) stat chart to the right
- "Predict your fire" button title
- Tableau slides link

"Using Neural Networking with One hot encoder results in the highest accuracy percentage which will give our model the best chance of quickly predicting a fire size. This will not only be helpful but also not require large amounts of time and technology to run. Thus providing an entertaining and informative tool for wildfire education."

Final Presentation Talking points:

- Fire occurrence is increasing over time as well as fire size
  - Reference tableau slide for chart of time vs fire size
- Fire cause is mainly man made-
  - use tableau to filter fire causes on map to show that human fires are more than natural fires
- Temp and wind speed are the two main factors involved for large fires
- Early detection is critical to prevent loss of life and property damage
- "The nation's federal wildland fire community is a large and complex organization across the **Department of the Interior's Bureau of Land Management, the National Park Service, U.S. Fish and Wildlife Service, and the Bureau of Indian Affairs; and the U.S. Department of Agriculture's Forest Service**."
- https://www.nifc.gov/

Notes from 2nd submission:
- ❖ Explain that it is not fire size over time and that we are predicting fire size IF a fire occurs.
- ❖ Over time is to support climate change argument that fires are increasing in quantity and size as the conditions begin to worsen.