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**](https://www.theguardian.com/us-news/2019/jun/21/us-plastic-recycling-landfills)**. 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)**

**Project Wildfire:**

**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**](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