WHY DOES MY CITY SMELL BAD

Syed Haider Karan Paresh Ramzi Adil

DS5110 Summer 2021

Summary:

The city of Portland has been collecting complaints from citizens regarding strong odors - in particular oil, that cause symptoms such as headaches. As shown in the article [1], there is a general public concern that the air quality may be causing health concerns and that may be tied to the oil storage and processing facilities that are located along the Fore River. In response, the local government of Portland has reached out to our class. They requested that our team, along with another team, investigate odor complaints.

We want to determine if there are any correlations between weather events such as wind direction or temperature as well as smell complaints from citizens of Portland, and geographical locations. These data sets include relevant information such as odor description, location of complaint, temperature, etc. The goals of the project are to coordinate the separate data sets into one frame and analyze odor complaints and weather events.

Proposed Plan:

Processing: Data preprocessing will need to be performed prior to any data analysis. The focus of our team is the geographical data so we will need to perform geographical specific pre-processing. In particular wind direction as measured in degrees will need to correspond to cardinal directions that can be used.

A major challenge we will face is in data preprocessing. Coordinating the weather data and smell data will be difficult in particular because of the different times the measurements were performed. At the lowest interval, weather data is measured every 1 minute[3], while odor complaints may occur just seconds after each other[2]. If we were to use larger time intervals the challenge is determining a metric for odors, because odor descriptions are not standardized, they vary greatly and cells can be empty.[2][4]

Preliminary results:

The results here show that we are able to load the datasets and perform the preprocessing and cleaning techniques by removing the “Nan" values and empty columns . The preliminary result showed an extra column for the wind direction which maps the degrees to the cardinal directions/points. This technique will provide clear data for mapping the further results.

```
df1.head() #weather data
```

	Time	Temp Avg	Temp Low	Temp High	Heat Index	Wind Chill	Temp (Day) Low
0	2021-07-18 00:00:00	68.0	68.0	68.0	68.0	68.0	68.0
1	2021-07-18 00:01:00	68.0	68.0	68.0	68.0	68.0	68.0
2	2021-07-18 00:02:00	67.9	67.9	68.0	67.9	67.9	67.9
3	2021-07-18 00:03:00	67.9	67.9	67.9	67.9	67.9	67.9
4	2021-07-18	67.9	67.9	67.9	67.9	67.9	67.9

```
df2.head() #Smell my city data
```

	epoch time	date & time	smell value
0	1624075697	06/19/2021 00:08:17 -04:00	5
1	1624090416	06/19/2021 04:13:36 -04:00	5
2	1624098728	06/19/2021 06:32:08 -04:00	4
3	1624119361	06/19/2021 12:16:01 -04:00	3
4	1624119890	06/19/2021 12:24:50 -04:00	1

	Wind Direction	Cardinal Direction
0	337	North
1	337	North
2	337	North
3	337	North
4	337	North
...	...	...
2791	270	West
2792	270	West
2793	265	West
2794	229	West
2795	225	South-West

2796 rows × 2 columns

Reference:

[1]“Portland, South Portland Residents Ask State to Do More about Smell from Large Oil Tanks.” *Mary Cate Mannion*, 21 Aug. 2019, www.wmtw.com/article/portland-south-portland-residents-ask-state-to-do-more-about-smell-from-large-oil-tanks/28777159.

datasets:

[2] https://smellmycity.org/data

[3] https://rainwise.net/weather/SMRO3

[4] https://seeclickfix.com/portland_2

[5] https://github.com/ds5110/stinky