**Project Report**

**A group of cars in a parking lot

Description automatically generated with low confidence**

**Data Visualization on**

**Uber Analysis : Impact of Weather on Number of Rides**

**Introduction**

Uber is a ride-sharing startup that has transformed the transportation sector by offering consumers a cutting-edge and practical mode of transportation. Uber has significantly impacted the traditional taxi sector and changed how people view transportation as one of the most important participants in the gig economy. Several studies have looked at Uber's effects on the transportation sector recently, including the growth of ride-sharing services and the demise of conventional taxi services.

With an emphasis on how the weather may impact ride demand, this data visualization project seeks to examine how Uber has affected the transportation sector. The study will examine the relationship between Uber's ride demand and meteorological variables like rain, snow, and temperature using data visualization tools.

The National Oceanic and Atmospheric Administration (NOAA) provided the meteorological data, while Kaggle provided the data for this project. The project will examine the patterns of demand for rides under various weather conditions as well as how the weather affects weekdays and weekends differently.

This project intends to offer useful insights into the impact of weather on Uber's trip demand during various weather conditions by studying these numerous aspects and research issues. Researchers and policymakers interested in the gig economy and the transportation sector, as well as Uber and other ride-sharing businesses, can benefit from the findings of this project.

**Research Questions:**

1. What is the relationship between the number of Uber rides and the weather conditions in New York Counties?
2. What are the top 5 zones that has highest average no of trips in various counties?

**Methodology:**

To determine how the weather affects Uber rides In New York City, our study included gathering data from a variety of sources. The main data source is Kaggle, which offers details on Uber journeys, including the pick-up time and date, the location ID, and another dataset of taxi zones to determine the pickup zone. I obtained information from the National Oceanic and Atmospheric Administration (NOAA) to examine the effects of the weather.

**Cleaning Data:**

1. Created date and hour to link the dataset.
2. Converted hourly visibility as below as per the data based on Beaufort wind scale Visibility scale.

|  |  |
| --- | --- |
| For visibility | Scale |
| 0-0.75 | Fog |
| 1-2.5 | Poor visibility |
| 2.5 - 5 | Moderate Visibility |
| 5 above | Good Visibility |

1. Hour present weather type variable categorized into Clear, fog, Light rain, moderate rain, heavy rain, Light snow, Moderate snow, and heavy snow based on the WEATHER PHENOMENA MATRIX.

<https://www.weather.gov/media/surface/WSOH8.pdf>

Reference link for Datasets:

[**https://www.ncei.noaa.gov/access/search/data-search/global-summary-of-the-month**](https://www.ncei.noaa.gov/access/search/data-search/global-summary-of-the-month)

[**https://www.kaggle.com/datasets/fivethirtyeight/uber-pickups-in-new-york-city**](https://www.kaggle.com/datasets/fivethirtyeight/uber-pickups-in-new-york-city)

**Analysis:**

1. **Total no of rides per county**

**A map of the united states

Description automatically generated with low confidence**

* From the above Map, we can say that Manhattan has a demand of 75.45% of the total uber rides in New York city.
* Staten Island has the least number of trips, 0.04% of the total uber rides.

1. **Total and Average trips per month by borough and zone.**

**A screenshot of a computer

Description automatically generated with medium confidence**

* The Average and total demand for uber rides for Manhattan borough in the month of February is high 59,081 and 16,54,279 respectively and Bronx brough in the month of January has least average and total rides of 728 and 22560 respectively.
* Overall, February has more demand for uber when compared to January.

1. **Average rides per weekday by Borough**

**A screen shot of a graph

Description automatically generated with medium confidence**

* From the above visualization, we can say Manhattan borough has huge demand on Friday and Saturday, for queens and Brooklyn borough there is high demand on Sunday and Saturday.
* Hence, we can say that Weekends have higher demand for uber than weekdays.

1. **No of rides per day by Month**

**A screenshot of a graph

Description automatically generated with medium confidence**

* **The above line chart is used to track the average rides per day by month.**
* **From Feb 18th to Feb 21st has more demand than rest of the days.**
* **From Feb 12th to Feb 15th the demand of uber was more due to Valentine’s Day being on Saturday and Monday is president day which is a long weekend.**
* **In case of January, Last week has more uber rides and followed by 3rd week of the month.**

1. **Average no of rides per day with respect to weather**

**A picture containing text, diagram, screenshot, plan

Description automatically generated**

* **The above line graph visualization is made using weather, windspeed, dewpoint and visibility as a parameter and can be filtered as required.**
* **Weather and Dew point has positive and negative correlation with no of rides.**
* **Wind is negatively correlated, and Visibility is positively correlated.**

1. **Top 5 zones with no of rides**

**A picture containing text, screenshot, line, plot

Description automatically generated**

* **The above bar graph represents the top 5 zones in each county.**
* **The top 5 Manhattan zones have more rides when compared to other zones.**
* **The Staten Island has no significant impact on no of rides in the new York city.**

1. **Avg no of riders per hour by visibility.**

**A picture containing text, screenshot, plot, diagram

Description automatically generated**

* **The above line graph shows avg no of rides per hour based on visibility.**
* **The visibility was categorized as fog which has no visibility, Poor, Moderate and Good visibility.**
* **From the above graph we can clearly say that impact of Visibility is high on no of riders with respect to Normal hours having visibility.**

1. **Percentage change in average rides with respect to Normal Weather**

**A screenshot of a graph

Description automatically generated with medium confidence**

* **Heavy rain has huge impact of 39.63% on the no of rides against the normal day.**
* **Moderate snow has negative impact of -13.11 % against the normal day.**

**Dashboard 1:**

**It provides a summary of rides per day/hour with respect to boroughs and zones and also effect of weather temperature, dew point, visibility and windspeed.**

**A screenshot of a computer

Description automatically generated with low confidence**

**Dashboard 2:**

**The below dashboard provides behavior for no of rides per day by month and top 5 zones with no of rides and the percentage change in average rides with respect to normal day.**

**A screenshot of a computer

Description automatically generated with medium confidence**

**Conclusions:**

**Weather:**

* **Heavy and moderate rain has a high impact of 39.3% and 21.05% concerning Normal weather.**
* **Weather and dew point have a mixed relationship with the number of rides.**
* **Visibility and Wind have negatively correlated with the average no of rides.**

**Location and Time:**

* Weekends starting from Friday, Saturday and Sunday were busier than normal days.
* Evenings after office hours starting from 17th to 20th hour the demand increases and generally we surge in pricing due to high demand.

**Further research questions:**

1. Adding prices and tourist places around the region can track the surge in pricing and verify the effect of cause for more no of rides for a particular zone.
2. Geo location can be added to track the density of rides.