

WEATHER IMPACT ON BIKE USAGE

Project Objective:

To analyze how weather conditions (e.g., temperature, precipitation, wind speed) affect bike-share usage patterns using historical bike-share data and weather data from an API.

Tools and Technologies:

1. **Programming Language:** Python
2. **Libraries:**
 - 2.1. Pandas, NumPy (data cleaning and analysis)
 - 2.2. requests (API calls)
 - 2.3. matplotlib, Plotly (visualization)
3. **APIs:**
 - 3.1. Open Weather Map API (or similar for weather data)
 - 3.2. Bike-Share Data Source: A public bike-share dataset
Chicago - <https://www.kaggle.com/datasets/nessada/divvy-tripdata-new>
 - 3.3. Open Weather Historical Bulk Downloadable CSV

Research Question:

"How do weather conditions impact bike-share usage patterns, including ride volume, duration, and station activity between Members and Casual Users in Central Chicago, and can current and forecasted weather data be used to predict usage?"

Supporting Questions:

1. How Does Weather Impact Ride Volume?
2. What Temporal Trends Emerge Across Weather Conditions?
3. How Do Different User Types Respond to Weather?
4. How Do Weather Variables Correlate with Overall Ridership?

Organize Project Files:

Create a project folder with the following structure:

```
bike-weather-analysis/
├── data/
│   ├── chicago_weather_data_2020_2023.csv      # Chicago Weather Data from 2020 to 2023
│   ├── 202004-divvy-tripdata.csv               # Bikeshare data for April 2020
│   └── ...                                       # Other monthly bikeshare datasets from March 2020 to November
2023
├── scripts/
│   ├── api_pull.py                             # Script for integrating the API and fetching weather data
│   └── data_cleaning.py                         # Script for cleaning and preprocessing bikeshare and
weather data
│   ├── exploratory_analysis.py                 # Script for exploratory data analysis and generating insights
│   └── config.py                               # Configuration file containing the API key for OpenWeatherMaps
├── output/
│   └── merged_weather_bike_data.csv             # Cleaned and merged dataset of weather and
bikeshare data
│   ├── visualizations/                         # Folder containing saved charts and figures from analysis
│   └── predictions.csv                         # File with predicted probabilities for bikeshare usage
├── documentation/
│   └── Weather_Impact_on_Bike_Usage_Project_Proposal_Plan_(Revised).pdf # Finalized project
proposal and plan
│   └── Bikeshare_Weather_Data_Analytical_Report.pdf # Analytical report based on the
project
│   └── Bikeshare_Weather_Data_Analytical_Presentation.pdf # Presentation summarizing
project findings
├── group_individual_work/
│   ├── scripts_Jitesh/                         # Folder for individual work and scripts by Jitesh
│   ├── scripts_Lovecy/                         # Folder for individual work and scripts by Lovecy
│   └── scripts_Sean/                           # Folder for individual work and scripts by Sean
└── README.md                                  # Project overview and instructions
```

Project Milestones and Timeline Plan:

Milestone 1: Project Ideation (Dec 3)

Finalize the project goal:

- "How do weather conditions impact bike-share usage patterns in Central Chicago, and can current and forecasted weather data be used to predict usage?"

Assign roles (e.g., data cleaning, analysis, API integration).

Milestone 2: Data Cleaning and Preparation (Dec 5)

Write the data_cleaning.py script:

- Load and preprocess both datasets (bike-share and weather).
 - o Remove duplicated and nulls (reduce to workable rows/columns)
 - Convert timestamps to datetime and align data by time.
 - Add derived columns like
 - o ride duration
 - o day of the week.
 - Save the cleaned dataset as merged_data.csv.
-

Milestone 3: Exploratory Data Analysis (EDA) (MVP) (Dec 7)

Write the exploratory_analysis.py script:

- Analyze usage patterns based on weather
 - o (e.g., ride counts by temperature or precipitation).
 - Create visualizations
 - o (e.g., line charts for trends, scatter plots for relationships).
 - Document key findings by section
 - o Save Findings Along with Chart for Presentation
 - o Save results and charts to output/.
-

Milestone 4: Regression Modeling (Dec 8 - 9)

Write the regression_model.py script:

- Perform linear regression using Python's scikit-learn.
- Predict bike-share usage (e.g., ride counts) based on weather variables like temperature and precipitation.
- Evaluate the model using simple metrics like R^2
- Save results and charts to output/.
-

Milestone 5: API Integration and Prediction (Dec 9)

Write the `api_prediction.py` script:

- Fetch real-time weather data using the weather API.
- Use the regression model to predict bike-share usage probabilities for current weather conditions.
- Output a CSV (`predictions.csv`) with the predictions and relevant metrics (e.g., confidence scores).

Example Prediction Output:

Timestamp	2024-12-10 14:00,
Location	Chicago
Weather	Cloudy
Temperature (C)	14 C
Busiest Station ID	86
Predicted_Rides	120
Predicted Length(hh:mm:ss)	00:03:24
Confidence_Score (predicted rides)	0.85

Milestone 6: Documentation & Presentation (Dec 10 – 11)

Write a comprehensive `README.md`:

- Include the research question, methodology, results, and API predictions.
- Add a guide for replicating the analysis.

Create slides with:

- Project objective and methodology.
- Key visualizations and insights.
- Explanation of regression results and predictions.
- Real-world application of predictions for stakeholders.

PRACTICE SLIDES

Milestone 7: Final Presentation (Dec 12)

- **Description:** Deliver the final presentation to the audience/stakeholders.
- **Due Date: December 12**
- **Tasks:**
 - o Present findings clearly and concisely.
 - o Be prepared for questions and feedback.

Timeline Summary

Milestone	Due Date
Project Ideation	December 3
Data Cleaning and Preparation	December 5
Exploratory Data Analysis (EDA)	December 7
Regression Modeling	December 8 - 9
Integration with Weather API	December 9
Documentation & Presentation	December 10 - 11
Final Presentation	December 12