VISHNU
UNIVERSAL LEARNING

Embedded Merge & Split Visual Adjustment of Data Grouping

Under the Guidance of

Guide Name: Mr.A.Rajashekar Reddy
Designation: Assistant Professor

Team: 5

17WH1A1216 - A.Supriya
17WH1A1220 - B.Divya Krupa
17WH1A1221 - BH.Deepika
17WH1A1237 - M.Sai Laya

Contents

Introduction

- Data is being produced enormously every day, which makes it challenging to manage and make sense.
- Retrieving results from extensive Data can be achieved using Data Visualisation, a part of Data Science that helps in revealing exciting patterns and observations of vast data.
- This project focuses on Data Visualisation by Embedded Merge and Split(EMS) that helps visualize the complex concepts of data.
- This system intends to visualize the data and group the data to match the users' requirements.

Abstract

- Data grouping is among the most frequently used operations in data visualization.
- It is the process through which relevant information is gathered, simplified, and expressed in summary form.
- EMS is a new interaction technique for direct adjustment of data grouping criteria.
- EMS technique can be designed to directly manipulate width and position in bar charts and histograms, of data grouping criteria.
- EMS can significantly reduce interaction time compared to WIMP-based technique.

Problem Statement

Our current system takes its roots from the interface **AvantGarde** (exploratory data analysis tool).

- To implement a interface with various panels in it depicting various visualizations.
- As part of this system, to visualize data into bar graphs, histograms, line graph, chore plot and matrix plot .
- To provide a panel to filter the data and columns that will help the user to visualize and extract patterns that are required for their analysis.

Literature Survey

AUTHOR	TITLE	DESCRIPTION	LIMITATIONS
Matthew N. O. Sadiku, Adebowale E. Shadare, Sarhan M. Musa and Cajetan M. Akujuobi	Data Visualization	Presenting data in graphical or pictorial form which makes the information easy to understand	The representations doesn't involve any user interactions
Inseok Ko, Hyejung Chang	Interactive Visualization of Healthcare Data Using Tableau	Procedure for the interactive visualization and analysis of healthcare data using Tableau as a business intelligence tool	supported using the Window/Icon/Menu/Pointer (WIMP) model which requires specific parameter adjustments over several steps.
Robert F. Eracher, Philip C. Chen, Jonathan C. Roberts, Craig M. Wittenbrink	Visual Data Exploration and Analysis VII	Easy to visually explore extremely large datasets. Loaded with both Infographics and Visual Explorer functionality	users constantly need to adjust the predefined data groupings created by visualization tools based on their evolving needs and interests.

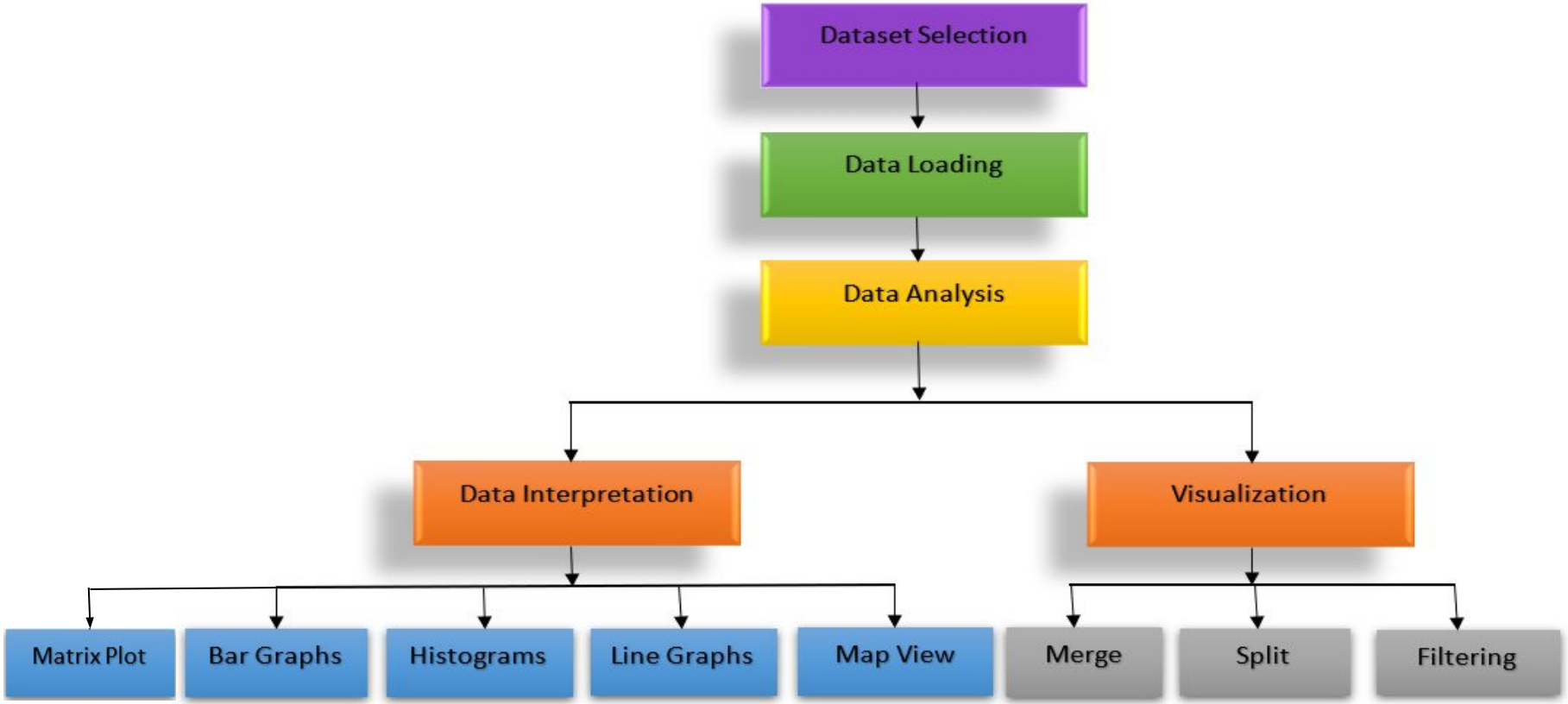
Existing Systems

- Many existing visualization tools such as Tableau and MS Excel support creation and presentation of grouped data.
- These tools perform data grouping and present the data and automatically performs data grouping based on statistical properties.
- But during visual data exploration processes, users constantly need to adjust the predefined data groupings created by visualization tools based on their evolving needs and interests.
- The WIMP (Windows, Icons, Menus, Pointers) technique can incur extra execution and cognitive costs especially as the number of available operations increases.

Proposed Methods

- The proposed system lets the user split the merged categorical and numeric data.
- By simple drag functionality, this interaction lets the user visualize the data.
- Various interaction methods like merge and split are proposed that reconfigures the graphs.
- Filtering option is proposed that updates all the modules accordingly.

System Architecture



Tools & Technologies

We used :

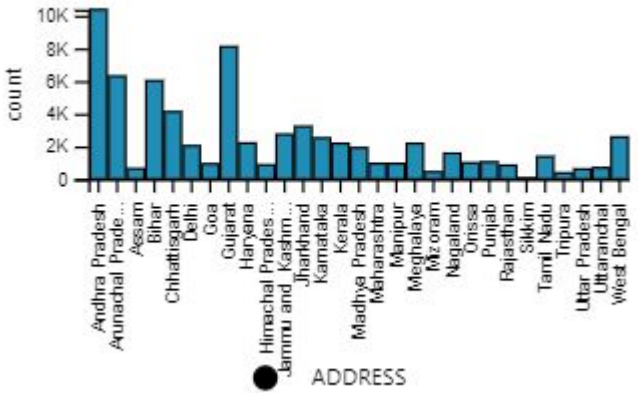
- D3.js (JavaScript library for manipulating data)
- Pycharm/ Visual Studio Code (IDE)
- Web Server (Apache)
- Web Standards
  - HTML (Hypertext Markup Language)
  - DOM (Document Object Model)
  - CSS (Cascading Style Sheets)
  - SVG (Scalable Vector Graphics)
  - JS (JavaScript)

Methods in D3.js

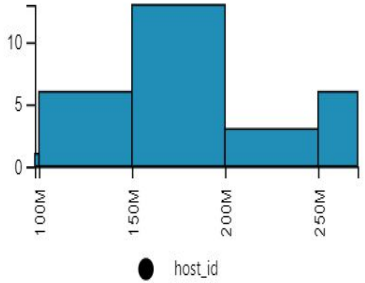
- d3.select("element")
- d3.selectAll(“element”)
- d3.nest()
- d3.json()
- d3.drag()
- d3.tip()
- d3.scaleLinear()
- d3.pie()
- d3.format(specifier)
- d3.interpolate(a,b)
- d3.mouse(“element)

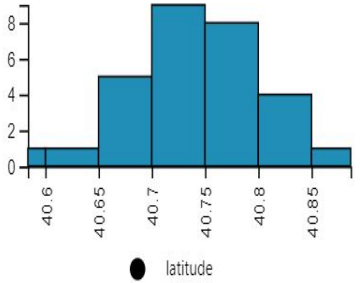
Project Modules

Bar graphs



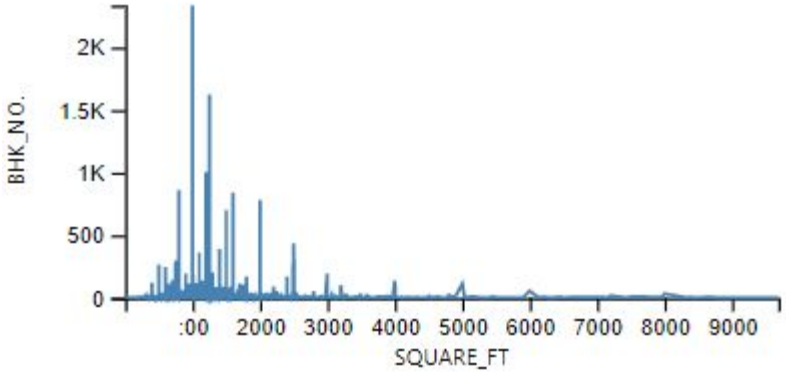
Histograms



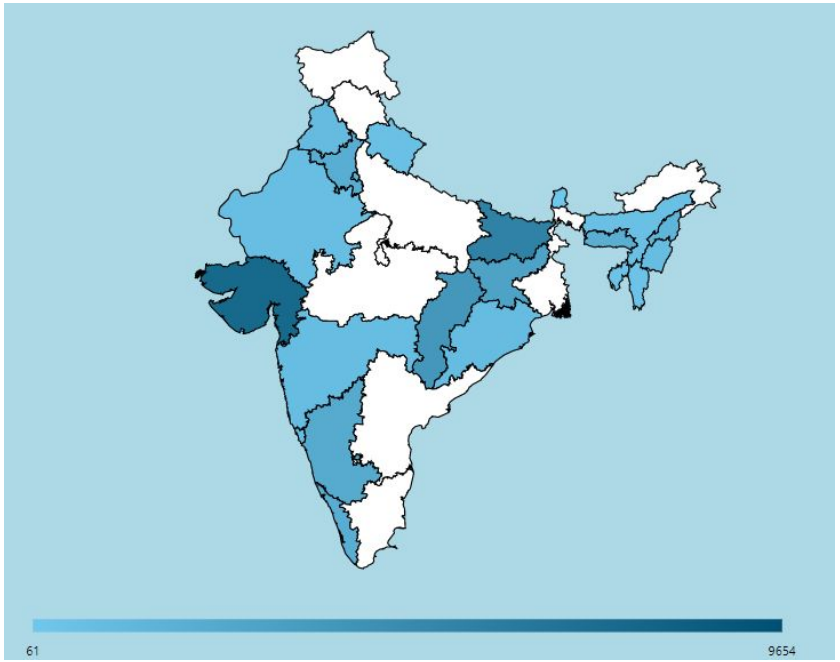


Project Modules

Line graphs



Map view



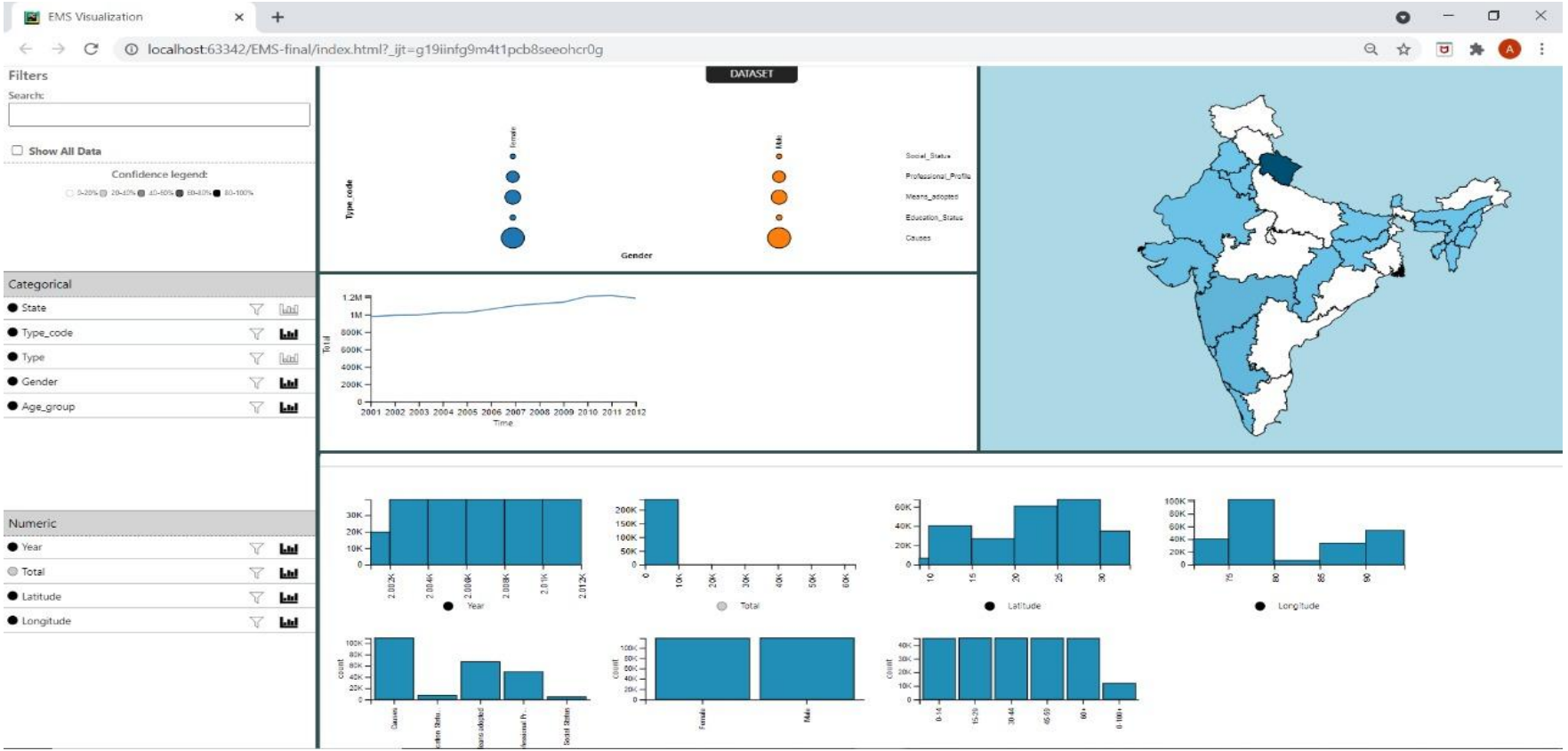
Project Modules

Matrix Plot

Output


EMS Visulization
localhost:63342/Embedded-Merge-and-Split-master/index.html?_ijt=3m5ba4p8lq1k3cu0ruparm4vek
Filters
Search:
Show All Data
Confidence legend:
0-20% 20-40% 40-60% 60-80% 80-100%
Categorical
name
host_name
neighbourhood_group
neighbourhood
room_type
last_review
Numeric
id
host_id
latitude
longitude
price
minimum_nights
number_of_reviews
reviews_per_month
calculated_host_listings_count
Brooklyn
Manhattan
Queens
Staten Island
Bronx
room type
Shared_room
Entire_home_or_apt
Private_room
neighbourhood_group
available_properties
Time
11541
486574
host_id
latitude
longitude
price
count
neighbourhood_group
room_type

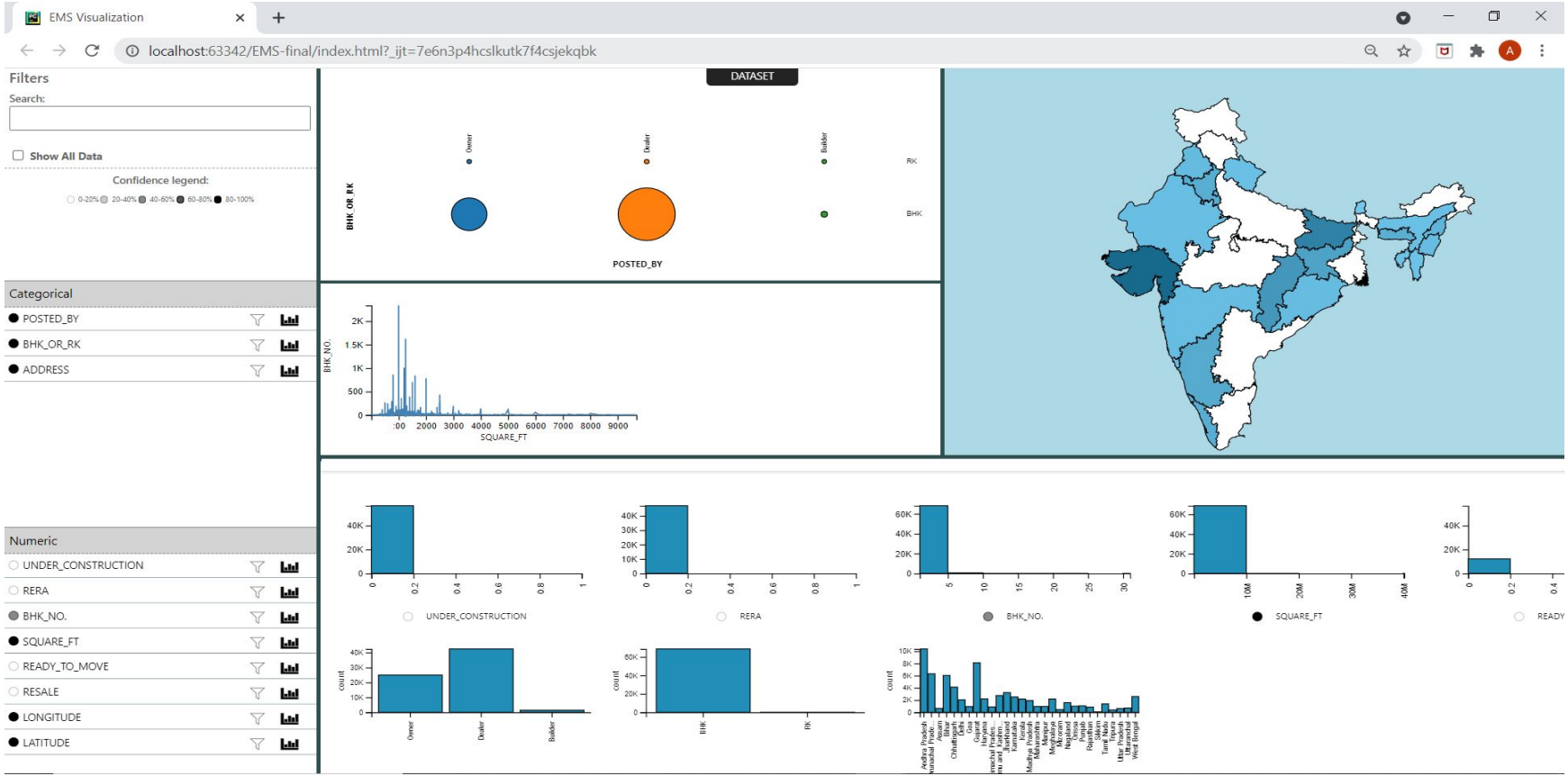
Output


EMS Visualization
localhost:63342/EMS-final/index.html?_ijt=g19iinfg9m4t1pcb8seeohcr0g
Filters
Search:
Show All Data
Confidence legend:
DATASET
Type_code
Female
Male
Gender
Social_Status
Professional_Profile
Means_adopted
Education_Status
Causes
Categorical
State
Type_code
Type
Gender
Age_group
Total
Time
Numeric
Year
Total
Latitude
Longitude
count

Output


EMS Visulization
localhost:63342/Embedded-Merge-and-Split-master/index.html?_ijt=3m5ba4p8lq1k3cu0ruparm4vek
Filters
Search:
Show All Data
Confidence legend:
0-20% 20-40% 40-60% 60-80% 80-100%
Categorical
Port Name
State
Border
Date
Measure
Location
Numeric
Port Code
Crossings
State
Measure
Crossings
Time
Port Code
Crossings
State
Border
Measure

Output

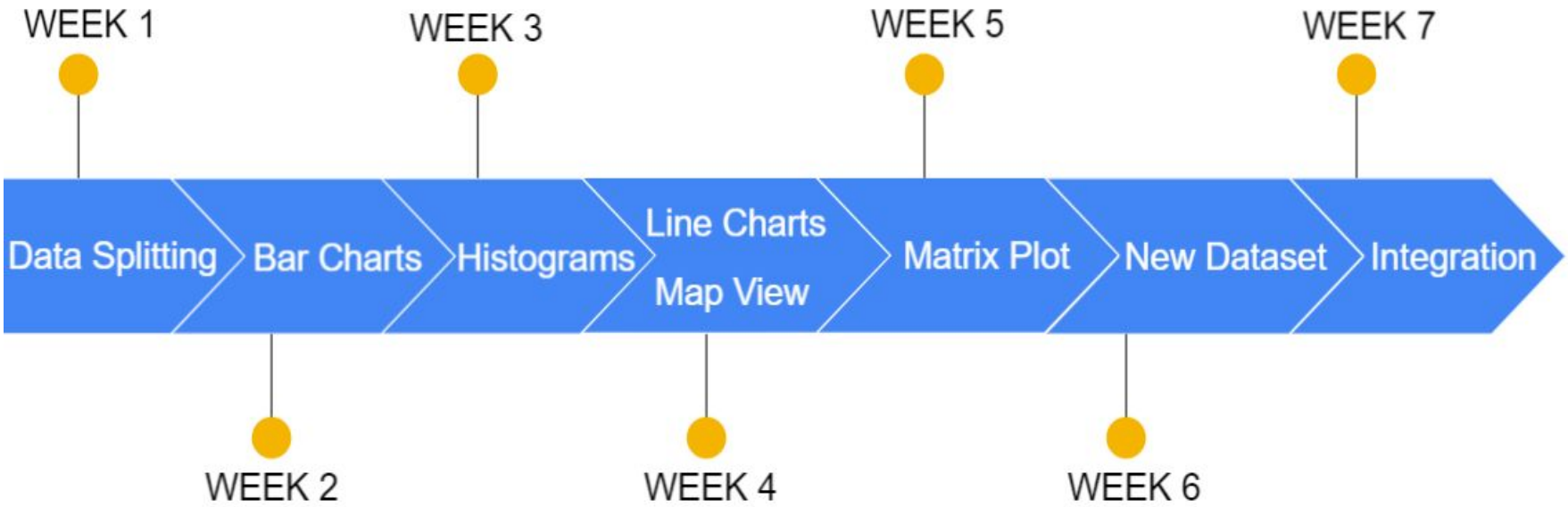

EMS Visualization
localhost:63342/EMS-final/index.html?_ijt=7e6n3p4hcslkutk7f4csjekqbk
Filters
Search:
Show All Data
Confidence legend:
0-20% 20-40% 40-60% 60-80% 80-100%
DATASET
BHK_OR_RK
Owner
Dealer
Builder
RK
BHK
POSTED_BY
BHK_NO.
SQUARE_FT
Categorical
POSTED_BY
BHK_OR_RK
ADDRESS
Numeric
UNDER_CONSTRUCTION
RERA
BHK_NO.
SQUARE_FT
READY_TO_MOVE
RESALE
LONGITUDE
LATITUDE
count

DEMO

Societal Impact

- ❏ Identify Trends Faster
- ❏ Identify Relationships More Effectively
- ❏ Simplify Complex Data
- ❏ Making Large Numbers Seem Relatable

Project Timeline / Schedule



Conclusion & Future Scope

- We introduced Embedded Merge & Split (EMS), a novel embedded interaction technique that enables users to adjust data grouping criteria by directly manipulating encodings used for presenting groups.
- We implemented EMS for bar charts, histograms, line graphs, map view and matrix plot for visualization of data.

Future Scope:

- We view this work as the first step towards exploring EMS.
- We want to explore and implement EMS techniques to other visualizations and data types.

References

1. D3js. Retrieved from https://d3js.org/
2. https://www.kaggle.com/dgomonov/new-york-city-airbnb-opendata
3. https://www.cc.gatech.edu/~aendert3/resources/Sarvghad2018Embedded.pdf
4. https://www.kaggle.com/akhilv11/border-crossing-entry-data/notebooks
5. https://www.kaggle.com/rajanand/suicides-in-india

6.https://www.kaggle.com/anmolkumar/house-price-prediction-challenge/tasks?taskId=2304

7.https://medium.com/multiple-views-visualization-researchexplained/what-is-visualization-research-what-should-it-be-8840a9ba658

Thank You