COMP47970 Visual Exploration Tool Design Document

Student Name: Aakash Sukre

Student Number: 23200858

#### Title:

US Post Offices

#### A map of the united states Description automatically generatedScreenshot:

#### Dataset overview:

This is a modification of a dataset provided sourced from the [TidyTuesday GitHub repository](https://github.com/rfordatascience/tidytuesday/blob/master/data/2021/2021-04-13/readme.md). The original dataset provides details such as post office names, locations (latitude and longitude), counties, establishment years, and discontinuation status. To prepare the dataset for visualization and analysis, I conducted several data manipulation operations using Python and relevant libraries. These operations encompassed tasks such as handling missing values, refining the dataset by removing irrelevant columns, and applying filters to include only relevant records based on specific criteria like stamp index, duration, and establishment year.

Additionally, I enhanced the dataset by adding a new column for the full state name derived from state abbreviations. Furthermore, I refined the dataset by excluding historical post offices established before 1910, focusing instead on more recent postal infrastructures. The culmination of these manipulations resulted in the creation of ['USA\_post\_offices.csv'](https://raw.githubusercontent.com/CLOUD007-stack/Information-Visualisation/main/USA_post_offices.csv), a refined and enriched dataset primed for comprehensive analysis and visualization.

*Design considerations:*

**Overall goal:** The overall goal of my visualization is to provide a comprehensive exploration of US post office data. Through interactive visual elements and dynamic data filtering, I aim to reveal insights into the distribution, characteristics, and historical evolution of post offices across different states and regions. This facilitates deeper understanding and analysis for historical research, urban planning, and geographical studies.

**State Vs Count of Post Offices Bar chart:** This bar chart shows the total number of post offices in each US state, sorted by the count of post offices per state. Sorting by this makes it easy to spot which state has the most post offices, but it can be tricky to find a specific state. Alternatively, sorting alphabetically would help with finding specific states but make it harder to see the highest and lowest counts. I have used a single steel blue colour for all states instead of different colours to keep things simple. When you select one or more state bars, the others fade to light grey. While I could have used grouped bar charts to show continued and discontinued data for each state, I decided to stick with one bar per state for a straightforward comparison of total post office counts across the US.

**Continuous status Vs Count of Post Offices Bar chart :** The bar chart visualizes the distribution of "True" and "False" values representing the operational status of post offices, with "True" indicating operational and "False" indicating discontinued. Initially, I experimented with using two buttons for visualization, but this approach lacked the ability to provide a clear comparison between true and false counts. The bar chart resolves this issue by offering a straightforward comparison through bar lengths, making it easier to identify individual true and false counts, especially those with higher frequencies. This visualization effectively highlights the distribution of post office statuses, enhancing understanding of operational trends and discontinuations across different regions or time periods.

**Stamp Index Vs Count of Post Offices Bar chart :** This bar chart illustrates the relationship between the stamp index values and the count of post offices. I initially attempted to visualize this relationship using a line chart but found that it did not provide a clear understanding. The advantage of the bar chart format is that it enables a straightforward comparison between different stamp index values and their corresponding counts of post offices. Each bar represents a specific stamp index value, with the height of the bar indicating the count of post offices associated with that stamp index. This bar chart offers a clearer insight into how stamp index values correlate with the number of post offices, facilitating the identification of any trends or patterns in post office distribution based on stamp index values.

**Dot plot map:** In this dot plot map, I visualize the geographic distribution of post offices across the United States, representing each state with a dot. This approach offers immediate insights into spatial patterns but may suffer from overplotting in densely populated areas. To enhance clarity, aggregating post offices into county-level summaries could be considered, though this would sacrifice state-level inspection facilitated by tooltips. Additionally, I utilise a TopoJSON map to highlight missing state data, providing context on dataset completeness. Integrated functionality dynamically updates the map based on selections made in other bar charts, enabling users to correlate post office distribution with stamp index or operational status.

**Interaction consideration:** In designing the interaction approach for this visualization, I have focused on cross-filtering as the primary method. This allows users to select subsets of the data in one chart, which then filters the data in another chart accordingly. For instance, in the bar chart depicting state vs count of post offices, users can select a specific state, which will dynamically filter the data presented dot plot map. Each chart offers selection based on different data attributes – the bar chart allows selection by state, while the stamp index vs count of post offices and continuous status (True/False) vs count of post offices charts enable selection based on state, stamp index, Continuous status. Moreover, the dot plot map features an interactive tooltip displaying all relevant information regarding post office establishment and discontinuation years, status, duration, and stamp index. By integrating these interactive elements, users can seamlessly explore the relationship between these attributes, such as identifying the locations of continued post offices and their corresponding stamp indices.