Based on your descriptions and the images provided, here's a detailed breakdown for each of your visualization systems:

**System A Description:**

**Visualization Components:**

1. **Dotted Line Plot (Top Left):** This plot presents the average speed (km/h) of taxi trips across different hours of the day. Each line represents one of the districts: Brooklyn, Manhattan, Queens, and Staten Island. The dots on the lines may facilitate identifying specific hourly averages per district.
2. **Heatmap (Bottom):** Displays the number of taxi rides against the day of the week (Y-axis) and hour of the day (X-axis). The color intensity within each cell reflects the count of taxi rides, with darker shades indicating higher volumes.
3. **Stacked Bar Plot (Right Side):** This visual represents the total number of taxi rides for each day of the week, broken down by district. Each color within a bar signifies a different district, helping to compare the distribution of taxi rides per district throughout the week.

**Interactivity:**

* Users can click and drag within the heatmap to select specific time intervals, affecting what is shown in the other two graphs.
* Similarly, users can select time intervals on the dotted line plot, and these selections will apply to the heatmap and the stacked bar plot.

**System B Description:**

**Visualization Components:**

1. **Faceted Area Plot (Left):** Displays the average speed of taxi rides across different hours of the day, with a separate plot for each district (Brooklyn, Manhattan, Queens, Staten Island). This faceted design allows for direct comparison of temporal trends between districts.
2. **Dot Graph (Top Right):** This plot uses the hour of the day (X-axis) and the day of the week (Y-axis) as its dimensions. Each cell in the grid hosts a circle whose size represents the number of taxi rides during that specific time and day.
3. **Pie Chart (Bottom Right):** Illustrates the total number of taxi rides segmented by each day of the week. Each segment's color corresponds to a different day, enabling a quick view of which days are busiest.

**Interactivity:**

* In the faceted area plot, users can click and drag horizontally across any area to select a time interval, which then filters the data across all plots.
* The dot plot allows for selection of specific hours and days by clicking and dragging across the relevant area.
* Clicking on a segment within the pie chart filters the data for that particular day of the week across all visualizations.

These systems each provide unique ways of exploring and understanding the taxi ride data, with interactivity designed to enhance user engagement and discovery.

**System C Description:**

**Visualization Components:**

1. **Bar Chart (Top Left):** Displays the average speed (km/h) of taxi rides for different districts (Brooklyn, Manhattan, Queens, Staten Island). This graph helps to compare the average taxi speeds across districts, providing insights into traffic conditions or possibly the nature of trips typically undertaken in each district.
2. **Stacked Area Plot (Top Right):** Represents the number of taxi rides for different districts from Monday to Sunday. This visualization demonstrates the distribution of taxi demand throughout the week, highlighting patterns of use and how they might vary by district.
3. **Strip Plot (Bottom):** Plots speed (km/h) of taxi rides on the Y-axis against the hour of the day on the X-axis. This plot can reveal the dynamics of taxi speeds throughout the day, indicating periods of high traffic or higher speeds due to lower traffic volumes.

**Interactivity:**

* In the strip plot, users can select a time interval horizontally, which could be useful for examining taxi speeds during specific parts of the day, such as rush hours or late at night.
* In the stacked area plot, users can select days horizontally, allowing them to filter the data to compare taxi usage across different days or to focus on specific patterns like weekend vs. weekday usage.

**1. The Data:**

**Title:** New York City Taxi Speed and Volume Analysis

**Description:** This dataset integrates records of taxi trips in New York City, detailing aspects such as speeds and trip counts across different districts and times. Derived from the a Kaggle competition dataset, the data encapsulate both temporal (hour of the day, day of the week) and spatial (district-based) dimensions, enabling a comprehensive analysis of urban transportation dynamics. Quantitative attributes include trip durations and distances, while categorical attributes encompass district labels and time classifications. This dataset supports multifaceted inquiries into urban mobility, revealing patterns of taxi usage in relation to time and geography.

**2. The Tasks:**

In exploring the dataset on taxi speeds and counts across different districts of New York City, users engage with three primary tasks: Analyse, Search, and Query. These tasks align with Munzner's framework in "Visualization Analysis & Design" and cater to a thorough investigation of the dynamic taxi usage patterns under various temporal and spatial contexts.

**Analyse:** Users aim to digest complex multi-dimensional data through visual representations. They can consume the data to uncover unknown trends, such as identifying specific hours or days with unusually high taxi speeds or volumes. Conversely, they might produce new insights, like annotating periods of significant traffic slowdown or surges in taxi usage. For instance, they could discover if a particular district consistently shows lower speeds during rush hours, indicative of potential congestion or if there are notable differences in taxi usage between weekdays and weekends.

**Search:** This involves pinpointing specific pieces of information within the visualization. Users can:

* **Lookup**: Directly find known targets, like the average speed of taxis in Manhattan at 8 AM.
* **Browse**: Explore data when the location is known but the target isn't, such as inspecting Saturday to see which hours are busiest.
* **Locate**: Seek out known targets with unknown locations, like identifying if there were any days with exceptionally high taxi speeds.
* **Explore**: Delve into the data without a specific target or location in mind, aiming to uncover patterns or anomalies across hours and districts.

**Query:** After locating data of interest, users perform deeper investigations:

* **Identify**: Extract comprehensive details from specific data points, like the exact number of taxis operating in Queens at a given hour.
* **Compare**: Analyze differences between data items, such as comparing the average taxi speed between weekdays and weekends or across different districts.
* **Summarise**: Generate a condensed overview of multiple data items, like the average number of taxis operating across all districts for each day of the week.

**4. Generalised Selection:**

For the taxi speed and volume analysis, a semantic structure is designed based on hierarchical data dimensions: Time (Day of Week, Hour of Day) and Space (Districts). Users traverse this structure starting from broader categories (e.g., entire weeks or all districts) to finer details (specific hours or districts). The traversal policy employs interactive visualizations allowing users to drill down into the dataset: selecting a district filters the temporal data views, while choosing a time frame updates spatial distribution metrics. This enhances user engagement by facilitating intuitive exploration and enabling comparative analyses across different dimensions.

**6. Design Comparison**

**Choice 1: Differentiating and Comparing Values for Different Districts**

**System A:** Utilizes a dotted line graph with separate lines in distinct colors to represent each district, offering clarity and simplicity in distinguishing between areas. The stacked bar chart further differentiates districts using various colors, making it easy to compare daily taxi ride counts side by side.

**System B:** Deploys separate area plots for each district to display the average speed versus hour, providing a clear and isolated view of each district’s performance over time without overlap, enhancing comparability at the expense of compactness.

**System C:** In the bar chart, different colors and separate boxes are used to differentiate between districts, providing a distinct visual segmentation. The stacked area plot employs varying colors to distinguish between districts within the same plot, enabling a cumulative comparison while maintaining individual district identities.

**Best Choice Analysis:** The choice depends on user needs: System A for users who prefer integrated temporal trends across districts, System B for detailed, non-overlapping district analysis, and System C for a balanced view that includes both segregated and integrated district data.

**Choice 2: Type of Graph for Hourly Changes**

**System A:** Employs a line plot in the dotted line graph to reflect hourly changes in average speed, which is suitable for tracking continuous data over time. The heatmap is used to represent hourly changes in taxi counts, providing a visual distinction between different times and frequencies.

**System B:** Uses an area plot for showing hourly changes, offering a clear visualization of volume and density over time which is beneficial for highlighting peak and off-peak periods.

**System C:** Implements a strip plot to depict hourly changes, providing a unique approach that combines aspects of line and bar graphs, showing discrete time points while still allowing for trend observation.

**Best Choice Analysis:** Each type serves a different purpose: line plots (System A) for clear trends, area plots (System B) for cumulative impacts, and strip plots (System C) for detailed yet concise time-point data visualization.

**Choice 3: Interactivity and Data Selection**

**System A:** Provides interactive selections through a heatmap and line graph, allowing users to filter based on time and day, applying selections across other visualizations.

**System B:** Facilitates interactive data exploration by allowing users to select time intervals in faceted area plots or specific times and days in the dot plot.

**System C:** Offers horizontal time selection in the strip plot and day selection in the stacked area plot, enabling focused data analysis based on user-specified criteria.

**Best Choice Analysis:** Interactivity level depends on intended user engagement; System A is for broad temporal selections, System B for detailed segment analysis, and System C for focused, user-driven exploration.

**Choice 4: Visual Encoding for Time and Date**

**System A:** Utilizes a linear time axis in the line plot and an ordinal axis in the heatmap, effectively encoding time in a manner conducive to trend analysis and daily activity patterns.

**System B:** Presents time as a continuous variable in faceted area plots and categorically in the dot graph, accommodating both continuous and discrete analysis.

**System C:** Adopts a uniform approach to time encoding in the strip plot and a categorical treatment in the stacked area plot, blending continuous and categorical perspectives.

**Best Choice Analysis:** Choice depends on data characteristics and user preference; continuous encoding (System B) for detailed analysis, categorical (System C) for broader trends, and a mix (System A) for versatility.

**Choice 5: Color Scheme and Accessibility**

**System A:** Employs distinct color schemes for differentiating districts and representing data density, balancing distinction with aesthetic appeal.

**System B:** Chooses a consistent color scheme across different plots, maintaining visual coherence and user focus.

**System C:** Adopts varying color schemes to highlight different data types and districts, ensuring clear differentiation while enhancing interpretability.

**Best Choice Analysis:** Consider accessibility and color blindness; System B’s consistent scheme is user-friendly, while System A’s and C’s diverse palettes cater to distinct separations but require careful color choice to avoid confusion.

**Choice 6: Data Aggregation and Presentation**

**System A:** Shows aggregated data in the stacked bar chart for overall trends, while offering detailed speeds in the line plot.

**System B:** Separates data by district with individual plots, providing clear, non-overlapping views for each area’s trends.

**System C:** Combines different data aggregation levels: detailed hourly speeds in the strip plot and cumulative weekly counts in the stacked area chart.

**Best Choice Analysis:** System B is ideal for users needing clear, segregated district data, System A for those seeking an overview plus detail, and System C for users requiring a balance between detailed and aggregated data insights.

Top of Form

Bottom of Form

In conclusion, each system has its strengths: A excels in integrated analysis, B in simplicity and segmentation, and C offers a balanced, comprehensive approach. The choice among them depends on user needs and analytical goals, influenced by the intricate trade-offs between complexity, detail, and usability, as discussed in the visualization principles and data encoding strategies from the course materials.