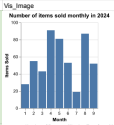
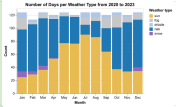
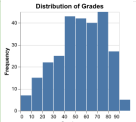
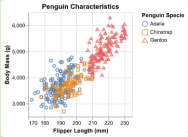
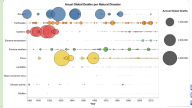
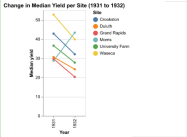
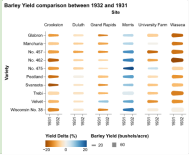
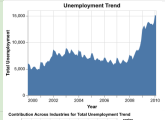
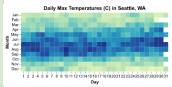
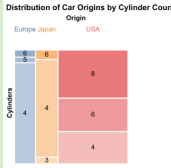
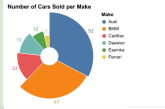
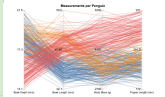
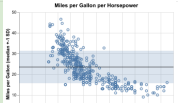
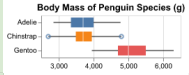
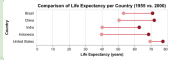
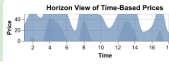
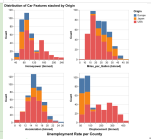
Vis Code	Visualization Type	Visualization	Vis URL	Identification Tasks	Comparison Tasks	Trend Analysis Tasks	Relationship and Correlation Tasks	Data Exploration 1	Data Exploration 2 (takeaway/main insight)	Possible Memory-Related question?
Vis_Code	Vis_Type	Vis_Image	Vis_URL	_ID	_C	_T	_RC	_DE1	_DE2	_MEM
BC	Simple Bar chart		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/simple_bar_chart.png	MC: Identify the month that corresponds to the maximum number of items sold.	MC: Compare the number of items sold for the months February (2) and August (8). Which one has a lower number of items sold?	TE: Describe the trend items sold in relation to the months of the year. Do the number of items sold increase, decrease, or remain inconsistent as the year progresses?	TE: Is there a discernible pattern or relationship between number of items sold and month? Describe any patterns you observe.	TE: Based on the chart, what insights can you infer about how the number of items sold correlates with the months in a year? Are there any notable trends or anomalies?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
SB	Stacked Bar Chart		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/stacked_bar_chart.png	MC: Identify the month that experienced the highest number of sunny days.	MC: Comparing April and October, in which month was the proportion of snowy days greater?	TE: How does the occurrence of rainy days trend throughout the years? Can you identify months where rain is significantly more frequent?	TE: Does the chart suggest a correlation between sunny days and rainy days within any particular month? Please describe any patterns observed.	TE: Upon reviewing the chart, highlight any months that stand out due to an unusually high or low frequency of foggy days.	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
DB	Diverging Stacked Bar Chart (with Neutral Parts)		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/diverging_stacked_bar_chart_(with_neutral_parts).png	MC: Which survey question garnered the highest level of agreement (combining Agree and Strongly Agree) from participants?	MC: Within Question 4's responses, which category—Strongly Disagree, Disagree, Neither, Agree, or Strongly Agree—constitutes the smallest percentage of the total responses?	MC: Analyzing the progression from Question 1 through Question 5, does the overall trend indicate that participants tend to become more agreeable or less agreeable as the survey advances?	TE: Examine the chart for a potential correlation between the proportion of agreeable responses (Agree and Strongly Agree) and disagreeable responses (Disagree and Strongly Disagree) across all questions. Do questions that have higher percentages of agreement also tend to have lower percentages of disagreement, or is there a more complex relationship evident between these response types?	TE: Based on the diverging stacked bar chart, how do response patterns of agreement and disagreement distribute across the different questions? Are there any notable trends in how participants agree or disagree with the content of the survey questions?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
HS	Histogram (from binned data)		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/histogram_(from_binned_data).png	MC: Identify the grade bin range that has the lowest frequency. What are the start and end values of this bin?	MC: Given the grades 55 and 85 within their respective bins, which one has a higher frequency in the histogram?	TE: How does the frequency of grades vary as you move from lower to higher grades?	TE: Explore the histogram to see if there's a consistent or varied relationship between grades and their associated frequencies. Does this histogram reveal any patterns in how students perform?	TE: Dive into the histogram and identify any anomalies or unusual patterns in the distribution of grades. Are there specific grades ranges that deviate significantly from the overall trend, perhaps indicating outliers, peaks, or gaps in the data distribution?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
HM	2D Histogram Heatmap		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/2d_histogram_heatmap.png	MC: Identify the bin with the highest count of records. What are the approximate IMDb and Rotten Tomatoes ratings for movies in this bin?	MC: Compare the concentration of records for movies with Rotten Tomatoes ratings above 80% to those with IMDb ratings above 8. Which rating system appears to be more generous, based on the density of high-rated movies?	TE: Is there a visible trend in the number of movies as you move from lower to higher IMDb ratings? Describe how the count of records changes across the heatmap.	TE: In the scatter plot created by the heatmap bins, do areas with higher counts of records suggest a strong relationship between IMDb and Rotten Tomatoes ratings?	TE: Explore the heatmap for any unusual patterns or anomalies. Are there any bins with unexpectedly high or low counts of records when considering the general trend of IMDb and Rotten Tomatoes ratings?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
CS	Colored Scatterplot		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/colored_scatterplot.png	MC: Which penguin species appears to have the longest flipper length on average, based on the scatterplot?	MC: Comparing the Adelie and Chinstrap species, which one tends to have a larger flipper length for a given body mass?	TE: Is there a visible trend in body mass as the flipper length increases across all penguin species? Describe the pattern you observe.	TE: How does the relationship between flipper length and body mass vary among the Gentoo, Adelie, and Chinstrap penguins? Is the correlation between these two measurements consistent across species?	TE: Explore the scatterplot for any unusual patterns or outliers within each penguin species group. Are there any individuals that significantly deviate from the general trends of their species in terms of body mass and flipper length?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
BU	Bubble Plot (Natural Disasters)		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/bubble_plot_(natural_disasters).png	MC: In which year and for which type of natural disaster was the largest bubble (indicating the highest number of annual global deaths) observed?	MC: Compare the annual global deaths caused by Earthquakes and Wildfires. Which disaster type has consistently led to larger bubbles (indicating higher death tolls) over the years?	TE: How has the size of the bubbles for Epidemics changed over the years? Does this suggest an increasing, decreasing, or stable trend in annual global deaths due to epidemics?	TE: Explore the relationship between time and disaster severity. Are there any discernible patterns indicating that certain decades or years are more prone to deadly natural disasters?	TE: Dive into the bubble plot and identify any outlier years where the death toll from a specific type of natural disaster significantly deviates from its usual pattern. What does this outlier suggest about the impact of that disaster type in that particular year?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
ML	Multi Series Line Chart		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/multi_series_line_chart.png	MC: Which company's stock achieved the highest price within the timeframe of 2000 to 2010, and in which year did this peak occur?	MC: Between Apple (AAPL) and Microsoft (MSFT), which company's stock showed a greater increase in price from the beginning to the end of the decade?	TE: Describe the overall price trend of IBM's stock over the decade. Did it experience more fluctuations compared to other tech stocks, or was it relatively stable?	TE: Examine the price trends of Google (GOOG) and Amazon (AMZN). Do these companies' stock prices show similar growth patterns or significant divergences during certain years? What might explain these observations?	TE: Investigate the line chart for periods of rapid growth or sharp declines in stock prices across all companies. Are there specific years where multiple companies experienced similar trends? Discuss these patterns and consider what factors might explain these movements.	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
SG	Slope Graph		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/slope_graph.png	MC: Which site experienced the highest increase in yield from 1931 to 1932?	MC: Compare the change in yield between Duluth and Morris from 1931 to 1932. Which site showed a greater improvement in their yield?	TE: Overall, did the majority of sites show an increase or decrease in yield from 1931 to 1932? Summarize the general trend observed across all the sites.	TE: Assess the slopes for any visible correlation between the initial yield in 1931 and the magnitude of change by 1932. Do sites with higher initial yields tend to show a greater increase or decrease in yield, or is there no clear pattern in how starting yields affect changes?	TE: Examine the graph for any outliers among the sites. Did any site experience an unexpected change in yield, significantly different from the others? Discuss any specific slopes that stand out from the general pattern observed.	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
CO	Comet Chart		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gemini/vegalite_examples/output/a_comet_chart_showing_changes_between_between_two_states.png	MC: Identify the barley variety and region combination that showed the largest increase in yield from 1931 to 1932 as indicated by the comet size and color on the chart.	MC: Compare the yield change of the 'Manchuria' barley variety in Crookston and Waseca. Which region showed less decline in yield from 1931 to 1932, based on the comet's size and color?	TE: Analyzing the comets across all regions, what trend can you identify for the 'Velvet' barley variety in terms of yield change from 1931 to 1932? Is there a consistent pattern of increase or decrease?	TE: Based on the comet chart, is there a noticeable pattern between the geographic location of the regions (Crookston, Duluth, Grand Rapids, Morris, University Farm, Waseca) and the yield change of barley varieties from 1931 to 1932? Do certain locations show consistently higher or lower yield changes, and how does this correlate with the variety of barley?	TE: Explore the comet chart and point out any interesting anomalies or patterns you observe in the yield changes for barley varieties across the different regions. Are there any varieties or regions that stand out for their performance change from 1931 to 1932?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?

Vis Code	Visualization Type	Visualization	Vis URL	Identification Tasks	Comparison Tasks	Trend Analysis Tasks	Relationship and Correlation Tasks	Data Exploration 1	Data Exploration 2 (takeaway/main insight)	Possible Memory-Related question?
AR	Area Chart		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/area_chart.png	MC: In which year does the chart show the highest amount of unemployment?	MC: Compare the levels of unemployment in the first half of the decade (2000-2005) with the second half (2005-2010). In which half does the chart indicate higher unemployment levels overall?	TE: Analyzing the trend from 2000 to 2010, how does the pattern of unemployment change over time? Is there a consistent increase, decrease, or does it fluctuate significantly within this period?	TE: Analyze the unemployment rate's fluctuation throughout the decade. Are there shorter cycles of rise and fall within the larger trend? How consistent are these cycles?	TE: Explore the area chart and identify any years or periods that show unexpected changes in unemployment levels. What might be some reasons for these changes, based on your observation or external knowledge?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
SA	Stacked Area Chart		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/stacked_area_chart.png	MC: In which year does the stacked area chart suggest the financial crisis likely occurred?	MC: Compare the change in unemployment for the Wholesale and Retail Trade and Construction industries over the decade. Which industry experienced a greater increase in unemployment from 2000 to 2010?	TE: Looking at the stacked area chart, which industry shows the most significant trend of increasing unemployment over the period from 2000 to 2010? Conversely, is there an industry that demonstrates a trend of decreasing unemployment?	TE: Explore the stacked area chart to identify which industries appeared to be most resilient to economic downturns and which were most vulnerable. What patterns can you discern about the impact of economic cycles on different sectors?	TE: Delve into the relationships between industries based on their unemployment trends. Are there any industries whose unemployment rates seem to move in tandem? This could suggest complementary or closely connected sectors.	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
HT	Annual Weather Heatmap		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/annual_weather_heatmap.png	MC: Which day and month recorded the highest daily maximum temperature in the year?	TE: How does the average daily maximum temperature in July compare to that in January in Seattle, Washington?	TE: Describe the trend in daily maximum temperatures from January through April. Is there a discernible pattern of temperature change as the year progresses?	TE: Can you determine any patterns or correlations between the time of year and the daily maximum temperatures in Seattle, Washington, based on the heatmap?	TE: Investigate the heatmap for insights on how the daily maximum temperatures fluctuate within a single month. Are there months with particularly erratic temperature changes, and what might this suggest about the weather patterns in Seattle, Washington?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
LS	Lasagna Plot (Dense Time Series Heatmap)		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/lasagna_plot_(dense_time-series_heatmap).png	MC: Which company showed the most significant increase in stock price over the ten-year period?	MC: Compare the stock price trends of AAPL and AMZN. Which company exhibited more volatility in its stock price over the observed period?	TE: What overall trend can you observe for IBM's stock price from 2000 to 2010?	TE: Analyze if there's a pattern between the time of year and significant changes in stock prices across the companies. Do certain years show a trend of price increases or decreases for all companies?	TE: Delve into the plot to find any periods of unexpected stock price behavior. Are there instances where a company's stock price deviated significantly from its general trend, and what might have caused these anomalies?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
MS	Mosaic Chart with Labels		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/mosaic_chart_with_labels.png	MC: Which country has the highest proportion of 8-cylinder cars as shown on the chart?	TE: How does the proportion of 4-cylinder cars from Japan compare to that from Europe?	TE: Identify any visible trends in the distribution of car cylinders across the three countries. Is there a noticeable preference for a certain number of cylinders in any country?	TE: Considering the varying widths of the columns for each country, what does this indicate about the overall distribution of cars by origin in the dataset?	TE: Examine the data for insights into how car engine configurations vary by the country of origin. Are certain configurations more prevalent in one country compared to others, and what might this suggest about automotive preferences or manufacturing trends in those regions?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
RD	Radial Plot		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/radial_plot.png	MC: Can you identify which car make had the highest sales according to the radial plot?	MC: Between Cadillac and Daewoo, which make shows greater sales numbers on the radial plot?	TE: Within the radial plot, do any car makes appear to dominate a particular range or section of the plot in terms of sales volume? What might this suggest about their market share or popularity?	TE: Based on the radial plot, can you discern any correlation between the sales numbers of the luxury car makes (e.g., BMW, Audi, Ferrari) compared to the non-luxury ones (e.g., Daewoo, Esemka)? Do luxury or non-luxury brands tend to have higher sales?	TE: Look at the distribution of car sales across the makes on the radial plot. Are there any unexpected leaders or underperformers in terms of sales volume? What does this suggest about consumer preferences or market competition?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
WF	Waterfall Chart of Monthly Profit and Loss		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/waterfall_chart_of_monthly_profit_and_loss.png	MC: Which month experienced the highest profit as illustrated by the plot?	MC: Compare the financial performance between the first quarter (January, February, March) and the fourth quarter (October, November, December). Which quarter demonstrates a higher overall profitability?	TE: What is the overall direction of the company's financial performance from the beginning to the end of the year? Is there a general upward or downward trend in profits or losses?	TE: Investigate if there's a link between the sequence of profitable and loss-making months. For example, do months showing profits tend to follow or precede those with losses, suggesting a cyclical pattern in the company's financial performance?	TE: Explore the data for evidence of seasonal impacts on the company's financial performance. Are there specific times of the year when the company tends to experience higher profits or losses?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
PL	Parallel Coordinate Plot		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/parallel_coordinate_plot.png	MC: Identify which penguin species tends to have the longest flipper length based on the plot.	MC: Compare the body mass of the Chinstrap penguin to that of the Gentoo penguin. Which species generally shows a greater body mass?	TE: Observe the patterns across all measurements for the Adelie penguin. Is there a noticeable trend in the physical characteristics that differentiates it from the other species?	TE: Analyze if there's a consistent relationship between beak length and beak depth across the three penguin species. Do longer beaks tend to be deeper, or is there no clear pattern?	TE: Explore the plot to determine which physical measurement (beak depth, beak length, body mass, or flipper length) most distinctly separates the three penguin species.	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
SD	Scatterplot with Mean and SD Overlay		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/scatterplot_with_mean_and_standard_deviation_overlay.png	MC: Determine the range of horsepower for vehicles clustered around the median MPG. What horsepower values are most commonly associated with this MPG range?	TE: Compare vehicles that fall within the MPG median and SD region to those outside of it. How does horsepower vary between these two groups?	TE: Describe the general trend between horsepower and MPG as visible on the scatter plot. Does an increase in horsepower correlate with a notable difference in MPG?	TE: Examine the scatterplot for patterns in how the spread of MPG values changes with increasing horsepower. Does a higher horsepower generally correspond to a wider or narrower range of MPG values, indicating variability in efficiency among more powerful vehicles?	TE: Examine vehicles that lie significantly above or below the median MPG. How does their horsepower compare to the rest of the dataset, and what insights does this offer about efficiency trends in vehicle manufacturing?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
BP	Box-Plot with Pre-Calculated Summaries		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/box_plot_with_pre-calculated_summaries.png	MC: Identify which penguin species shows the greatest median body mass based on the box plot.	MC: Compare the range of body mass for the Gentoo penguin against that of the Adelie penguin. Which species demonstrates a greater variability in body mass?	TE: Describe any trends or patterns you observe in the body mass distribution among the three penguin species. Is there a clear difference in body mass that distinguishes one species from the others?	TE: Analyze the relationship between the species of penguin and their body mass variability. Does a certain species tend to have more outliers in body mass, indicating exceptional individuals that deviate significantly from the species' median body mass?	TE: Focus on the interquartile ranges (IQR) and outliers in the box plots for each penguin species. What does the spread and presence of outliers tell you about the variability in body mass within each species? Could environmental factors or measurement errors play a role?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
DP	Ranged Dot Plot		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/ranged_dot_plot.png	MC: Which country has the highest life expectancy in 2000 among Brazil, China, India, Indonesia, and the United States, as shown on the ranged dot plot?	MC: Compare the increase in life expectancy from 1955 to 2000 between China and Indonesia. Which country saw a greater improvement?	TE: Identify the general trend in life expectancy changes from 1955 to 2000 across all included countries. Is there a consistent pattern of increase?	TE: Examine the relationship between the initial life expectancy in 1955 and the rate of increase by 2000. Do countries starting with a lower life expectancy in 1955 tend to have a larger or smaller increase by 2000?	TE: Investigate the plot for any anomalies or outliers in the data. Are there countries whose life expectancy either increased by an unexpectedly large amount or showed little to no improvement? What might account for these discrepancies?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
HR	Horizon Graph		https://raw.githubusercontent.com/cnobre/vis-repository/vis_generation/gpervis/vegalite_examples/output/horizon_graph.png	MC: Can you identify the point in time where the price reached its lowest point on the horizon graph?	MC: Compare the prices at time 4 and 17. Which time has a higher price?	TE: How does the price trend in the first half of the time period displayed on the horizon graph compare to the second half? Is there a notable change in volatility or direction?	TE: Examine the time elapsed between peaks in price. Does there appear to be a consistent or unconsistent relationship between the time between peaks in price?	TE: Upon examining the horizon graph, can you identify any cyclical patterns or regular intervals at which prices peak or dip? How consistent are these cycles over the entire time period displayed, and what might this suggest about the underlying factors influencing price fluctuations?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?

Vis Code	Visualization Type	Visualization	Vis URL	Identification Tasks	Comparison Tasks	Trend Analysis Tasks	Relationship and Correlation Tasks	Data Exploration 1	Data Exploration 2 (takeaway/main insight)	Possible Memory-Related question?
RC	Horizontally Repeated Charts		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gen/vis/vegalite_examples/outputs/horizontally_repeated_charts.png	MC: Identify which origin of cars generally has the highest Horsepower.	MC: Compare the Miles per Gallon distributions in the histogram: do American cars tend to have higher miles per gallon than Japanese cars?	TE: From the histograms, can you determine if there's a general trend in acceleration across cars from Japan, Europe, and the USA? Is there a noticeable difference in acceleration patterns?	TE: Examine the histograms for MPG, Horsepower, Acceleration, and Displacement. Can you identify a relationship between the MPG and Horsepower histograms that suggests how the origin of a car (Japan, Europe, USA) influences its fuel efficiency and power?	TE: Based on the distribution patterns seen in the histograms, what strengths and weaknesses can you deduce about cars from each origin (Japan, Europe, USA) in terms of MPG, Horsepower, Acceleration, and Displacement?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
CH	Chloropleth of Unemployment Rate per County		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gen/vis/vegalite_examples/outputs/choropleth_of_unemployment_rate_per_county.png	MC: Identify the state containing the county with the highest unemployment rate on the map.	MC: Compare the unemployment rates of coastal vs. inland counties. Based on the color coding, which area generally shows higher unemployment rates?	TE: Are there any visible regional trends in unemployment rates across the map? For instance, do certain states or geographical regions exhibit consistently higher or lower rates than others?	TE: Do you observe isolated counties with significantly higher unemployment rates compared to their surrounding areas, or are there broader regional trends where contiguous counties share similar unemployment rates?	TE: Explore the choropleth map for any emerging patterns or anomalies in unemployment rates. Are there unexpected clusters of high or low unemployment rates that stand out?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?
OD	One Dot per Zipcode in the US		https://raw.githubusercontent.com/cnobra/vis-repository/vis_generation/gen/vis/vegalite_examples/outputs/one_dot_per_zipcode_in_the_u.s.png	MC: Identify the starting digit of zip codes predominantly found in California as indicated by the color on the map.	MC: On the map, do the zip codes on the East Coast generally start with higher or lower digit numbers compared to those on the West Coast?	TE: Examine the correlation between the starting digits of zip codes and their geographical regions. How does the map illustrate the organization of zip code assignment across the United States?	TE: Examine the relationship between the geographical location and the density of zip codes represented by the colored dots on the map. What patterns can you discern about how zip code density varies across different regions?	TE: Look for patterns in the number of zip codes in the East vs. West coast. Does the allocation of zip codes vary between these two areas? If so, how?	TE: What would you say is the main takeaway/insight from this chart?	TE: What did you learn?