Design of Visualizations for "Startup Funding in California" Project Write-Up

Website Link: https://arakshe.github.io/

Stacked Bar Chart

This horizontal stacked bar chart was created using Altair. It visualizes the total startup funding across different markets, broken down by funding round, to give both a macro and micro view of the venture capital landscape.

Layout & Chart Type:

A horizontal stacked bar chart was used to improve readability of long category labels (market names). This choice ensures the visual is clear and well-structured, supporting a clean and professional layout as required.

Color Customization:

The bars are color-coded by funding round, using a custom sequential color palette ranging from light to dark blue to represent early to late-stage funding rounds (from "seed" to "round_h").

Labeling & Legends:

There are labels in the plot, with the x-axis as Total Funding (USD) and Y-axis as Market. Also on the legend, you can see the funding rounds, which are colored from light blue to dark blue as the funding rounds progress to advanced stages.

Interactivity:

An interactive dropdown filter is implemented at the bottom of the chart, allowing users to filter by a specific funding round to compare which industries get more funding per rounds. Or the user can view all rounds.

Sorting Logic:

The markets are sorted in descending order based on total funding, helping highlight the top-funded sectors at a glance, in line with the expectation that visualizations follow a logical, data-driven structure.

Consistency Across Website:

The font choices, color palette, and plot layout are consistent with the theme of the entire project website.

By these design choices. This visualization reveals that Software, Biotechnology, and Enterprise Software dominate in overall funding. These markets also show a diverse distribution across funding stages, indicating sustained investor interest throughout startup lifecycles. In contrast, markets like File Sharing, Communities, and Big Data received significantly less funding and were mostly concentrated in earlier rounds, suggesting limited scaling or reduced market traction. This insight contributes to our broader project narrative by illustrating where venture capital is concentrated and how it flows across industries and funding stages.

This stacked bar chart shows how startup funding is distributed across different markets and funding rounds. The markets are sorted by total funding, showing that Software, Biotechnology, and Enterprise Software are the top-funded sectors, each receiving billions of dollars in investments. These top 3 markets show a healthy mix of funding across early to late rounds, suggesting more growth potential and continued investor interest. In contrast, markets like File Sharing, Communities, and Big Data, indicate relatively limited funding. The color gradient illustrates how funding shifts through each stage, from seed rounds (light blue) to later stages like round_f, g, and h (dark blue). Many markets end after round_e or have little to no funding in round_g and round_h, suggesting that only a select group of companies survive or require funding in those later stages. The chart also shows variation in how funding is allocated within each industry. For example, Analytics has a large portion concentrated in round_f, possibly reflecting scaling efforts. With the dropdown filter, users can select specific rounds.

Treemap

This treemap was created using Plotly. The treemap was the best to visualizes the distribution of funding types (e.g., venture, private_equity, seed, etc.) grouped by the number of funding rounds startups received. These are broken into three categories: 1–3 Rounds, 4–6 Rounds, and 7+ Rounds. In the treemap, you can clearly see how the funding rounds are grouped under the categories, and which rounds are more relevant. This visualization shows how funding strategies shift as companies mature and raise more capital.

Layout & Chart Type:

A Treemap was chosen for its ability to represent hierarchical and proportional relationships in a space-efficient way. It allows viewers to quickly see which funding types dominate within each round category. Each parent group (1–3 Rounds, 4–6 Rounds, 7+ Rounds) contains sub-rectangles representing different funding types, with area size proportional to total funding.

Color Customization:

Each of the three round categories is visually separated by custom group coloring:

- Green for 1–3 Rounds
- Blue for 4–6 Rounds
- Orange for 7+ Rounds

This use of different colors helps the viewer clearly distinguish between funding stages and satisfies the requirement for customized styling.

Labeling & Legends:

Labels are directly embedded into the treemap blocks, showing funding types.
The hierarchical structure also shows the grouping by round count.
No axis is used, which suits the nature of this chart type and keeps the layout clean and focused.

Interactivity:

The chart includes interactive tooltips that appear on hover, displaying detailed funding information per block (such as funding type and value). This adds value beyond a static visualization and helps the user understand both categorical grouping and funding volume. This fulfills one of the required interactive visualization elements.

Sorting Logic & Structure:

Funding types are automatically sorted within each group by total amount, with the largest types (e.g., venture capital) appearing more prominently. This aligns with the requirement for data-driven structure and clarity.

Takeaway:

This treemap shows that venture funding dominates in all startup stages, especially within the 1–6 rounds range. As companies move into the 7+ funding round stage, there's a noticeable increase in the role of private equity and debt financing, indicating a shift toward later-stage, more complex funding mechanisms. The presence of seed and angel investments is minimal in later stages, which makes sense given their typical early-stage usage. This plot contributes to the overall narrative by helping us understand how the type of funding evolves as startups scale.

This treemap visualizes how funding types differ across startups based on the number of rounds they have raised. Venture funding dominates in all three stages. In all 3 categories(1–3, 4–6, and 7+ rounds), Venture funding dominates, which shows its key role in startup financing regardless of maturity or whether they have progressed more rounds. Venture funding is a little less as a company progresses to later stages because However, as startups progress to later stages (7+ rounds), there is a noticeable increase in the proportion of private equity and debt financing, suggesting a shift toward more complex or institutional funding sources. In contrast, early-stage companies rely more heavily on seed and angel investments, which shrink significantly in the later stages. This visualization highlights the evolving financial strategies of startups as they scale and mature.

This treemap reveals how funding types shift as startups progress through different stages of fundraising. Venture capital is the dominant funding source in the early and middle stages, particularly within the 1–3 and 4–6 rounds groups. However, in the 7+ rounds category, while venture capital is still present, its overall volume declines compared to earlier stages. This can be due to several factors. First, fewer startups reach such late stages, naturally lowering the total funding volume. Second, as companies mature, they gain access to other financing options such as private equity or debt financing that can either supplement or replace venture capital. As companies move forward in the process, there is a shift towards more complex and institutional funding sources. Also, venture capital firms often aim to exit their investments before or around this stage, especially if the company is heading toward an IPO (when a company goes public) or acquisition(when a company is acquired by another firm). Finally, mature companies may prefer funding options that don't require giving up ownership, such as debt, since they may already be

generating revenue. 7+ rounds group reflects a more diversified and strategic funding structure, while early-stage companies rely on seed and angel investments, which shrink towards the later stages.

Box Plot

This enhanced box plot was developed using D3.js, a powerful JavaScript library for data-driven document manipulation, to present a dynamic and insightful visualization of startup funding distributions across various investment rounds (A through G).

Purpose & Insight:
The plot captures both central tendency and variability in funding amounts across rounds, showcasing how investments evolve from early to later stages. The box plot format helps highlight medians, interquartile ranges, and extreme outliers—all crucial for understanding investor behavior and market maturity.

Regression Line:
To uncover overall trends in median funding amounts, I incorporated a regression line (in dark red) drawn through the median values of each round. This line uses a monotonic curve to reflect the general trajectory of funding. Its slight drop at later stages (F and G) hints at potential saturation or selectivity in large funding rounds—a subtle but significant insight.

Design Choices:
Clean Aesthetic: The plot maintains a minimalist design with a white background and clear labeling to enhance readability.
Interactive Elements: Tooltips are integrated to allow users to hover over elements and view detailed stats such as median values, exact funding amounts, and data spread without cluttering the interface.

Color Strategy:
Outliers are emphasized in orange-red, while box plots retain a professional blue tone, drawing attention to statistical anomalies while maintaining visual harmony.
Logical Ordering: Funding rounds are arranged in sequential order (A to G), creating a natural narrative from early to late-stage funding patterns.

This visualization was carefully crafted to blend aesthetics, interactivity, and analytical depth, making it suitable for both technical audiences and business stakeholders interested in venture funding dynamics.

Interactive Map

The interactive “Map of Startups in the Bar Area” has two layers to it. One of the layers is the bubble markers, and the other layer is a heatmap. We chose to include the bubbles to highlight the average funding amount in a certain location shown as the hue of the blue color, and then the

size of the bubble to represent the count of the number of start-ups in that area. We also decided to include an intensity heatmap as another layer to show the density of startups spread across different areas in California's Bay Area. There is an option to on/off each of the two layers on the bottom right side of the map –the viewer can see both together or each one separately. The option to view or not view the layers is an interactive technique we used here so the viewer has a choice of what they want to interpret in the graph based on what they want to find out. For the bubble chart, an entrepreneur who is looking to raise funding would look at the lightness or darkness of the color of the bubble to understand which areas within the Bay Area may have a higher likelihood of getting access to funding. The lighter blue color represents lesser average funding and the darker blue represents a higher level of average funding. We made sure to select and use only one color with the light-to-dark hue because it intuitively makes more sense to understand the progression of funding, rather than having multiple different colors. The average funding, which is in USD (based on the currency of the values in the dataset), ranges from a low of $827 to $66.3M. Furthermore, the size of the bubbles represents the startup count at a certain location, so the larger the bubble, the larger the number of start-ups in that area. The legend on the upper right-hand side of the plot also includes the range of the bubble sizes, which we ended up applying logarithmic transformation in the code to enhance the readability of the plot and sure the bubbles are proportionate. The zoom-in and out feature on the top left corner of the map can help the viewer better see the smaller bubbles find multiple bubbles/locations and hover over the bubbles to gain more information about the area. We also included the interactive tooltip feature so that when the viewer hovers over a bubble, they can see different relevant statistics of start-ups in that particular location, including the number of start-ups, the average funding amount, the top market, and the top-funded start-up. The top market is helpful in understanding which area may be better for a certain start-up idea, as a lot of investors generally like to specialize their investments within particular industries. Top-funded start-up is another relevant statistic as viewers can understand if there is any bias involved in regards to whether a particular area is dominated by the operations of one or a few particular start-ups (monopoly). Moving on, the heatmap intensity layer intuitively captures the big picture of the density of start-ups around the Bay Area and refers to the legend which is on the bottom left side of the map, blue represents a lower density where green is medium density, and red is high density. This allows the viewer to quickly grasp an understanding of the density of start-ups around the Bay Area.

Heatmap

The heatmap, which is created using Altair, shows the total funding given based on the first initial of the name of a start-up against the market that the start-up belongs to. We decided to use the top 30 most trending markets based on the dataset, which also aligns with how the top 30 markets were chosen to be represented in the stacked bar chart as well because most viewers are interested in these markets. The names of the top 30 markets are ordered alphabetically, from top to bottom. The different letters from A-Z are shown on the x-axis, which represents the first letter initial of the start-up names. This heatmap is created to understand if any branding trends budding entrepreneurs can be aware and informed of when building their start-up, and if there are particular types of start-up names that achieve higher credibility and brand loyalty and

attract higher funding in certain industries. The legend on the top right-hand side of the plot represents the funding amount, which ranges from $0 to $2,000,000,000, and the lesser the amount of funding the lighter the shade of the blue color. Only one color is chosen to be represented, blue, because it is more intuitively easy for the viewer to understand lower to higher funding amounts when trying to interpret the heatmap at a glance. The markets on the y-axis instead of being on the x-axis also significantly enhance the readability of the plot. Adding on, we also added the tooltip feature so that is easier to read the startup initial, market, and funding amount (in USD) when looking at one particular box out of many boxes on the heatmap and enhance visibility and readability overall.

Line Plot

This line plot emphasizes key financial events through vertical dashed reference lines and labels, which provide immediate historical context for changes in funding. The clean layout, with a single color representing the primary data line, ensures that viewers can quickly grasp the overall trend without visual clutter. The x-axis clearly displays the funding year, while the y-axis is scaled appropriately to emphasize both large spikes and subtle shifts in investment levels. The vertical guides, paired with concise event labels, help pinpoint the moments in time where external factors influenced startup funding trajectories.

To enrich user engagement, interactive features allow viewers to hover over data points and see specific funding values and event details, offering more in-depth insights without overcrowding the main chart. These tooltips or pop-ups also let users connect particular funding fluctuations directly to the labeled events, making it simple to explore correlations between significant global or market happenings and startup funding behavior.

Bar Plot by Market

This horizontal interactive bar plot by industry cluster, which is created using Altair, captures the top 10 markets within a certain industry cluster and their respective funding amounts (in USD). This data can help entrepreneurs better understand the scope of their product or service idea within a particular industry, based on the amount of funding that is typically allocated to that type of product or service. There is a dropdown interactive feature on the plot with the option to choose between 11 different industry clusters. For the top 10 markets, we decided that it is good and sufficient enough to give enough insight within a particular industry cluster. Once the viewer changes the industry cluster from the drop-down, we made the bar plot so that it changes the bar plot to show the top 10 markets and the total funding based on the cluster selected so the viewer can access more focused and relevant data. We also incorporated a tooltip interactive feature that shows the market, total funding ($), and cluster for better readability. Additionally, we decided to keep the color of all the bars the same to maintain consistency and clarity.