Spring Updated Project Plan

Hammon Vieira Dutra

Marcel Alves Pereira

Tiago dos Santos

Van Anh Nguyen

Yua Tsuchihashi

Minerva University

FE51: Spring Updated Project Plan

February 8, 2025

Team Facts

Team name: Data Hunters

Original question: How might we transform large publicly available data sets into engaging journalism that holds public entities accountable?

Word count: 23

Your Civic Partner

Our Civic Partner, Emma Steifel, is a Minerva alumna who graduated in Data Science and Arts & Humanities. As a data journalist at the San Francisco Chronicle, Emma writes for a city-scaled publication known for its evidence-based coverage of local issues. Emma helped us define a better project to focus on by giving insights about how our previous ideas were beyond our scope, requiring too much ability, effort, and time, and how our projects weren't innovative, since data analysis tools are abundant and journals already have a team focused on analyzing engagement. In our most recent meeting with Emma, she helped us define the analysis model we will be using. We were considering a chronological or spatial analysis, and since we are trying to compare the past with the present situation, she advised us to do a chronological analysis.

Word count: 138

Latest Project Direction: Refined Challenge Question and Proposed Spring Deliverable

Refined Research Question

How can we employ data analysis from San Francisco's previous tech boom to create an engaging journalistic piece that explains housing availability trends influenced by the recent AI

boom while documenting our methodology to provide a guideline for creating data-driven articles?

Proposed Spring Deliverable

Our team aims to develop a **data-driven news article** exploring how the AI boom in San Francisco affects the housing crisis across different neighborhoods. Additionally, we will document our research and analysis process in a separate guide to help non-experienced students replicate our approach to data-driven journalism.

Analysis Framework

Descriptive Analysis: Historical Context

The goal is to understand housing availability and pricing trends during the previous tech boom. To achieve this, the necessary data includes historical rental prices, property values, housing availability, and a timeline of the tech boom (e.g., company growth, and industry peaks). The analysis involves calculating percentage changes in housing prices and availability, visualizing trends using time-series graphs, and identifying the most impacted neighborhoods.

Comparative Analysis: Tech Boom vs. AI Boom

This section aims to compare housing market trends from the previous tech boom with the current AI boom. The data required includes recent housing data and AI industry indicators such as AI startup growth, VC funding, and job postings. The analysis focuses on mapping geographic trends to highlight affected areas and examining gentrification and displacement patterns in both booms.

Predictive Analysis: Emerging Patterns

The objective is to forecast housing trends influenced by the AI boom. This requires recent housing and AI industry data, along with economic indicators such as income levels, migration patterns, and construction rates. The analysis involves developing models to predict changes in housing prices and availability. Further research is needed to explore the use of machine learning in identifying future hotspots for housing shortages or price surges.

Spatial Analysis: Neighborhood Impacts

This section examines how specific neighborhoods are geographically affected by the AI boom. The necessary data includes neighborhood-level housing prices and inventory, as well as AI industry concentration by location. The analysis involves creating heat maps of housing affordability and availability and overlaying AI company growth data with housing trends to identify spatial correlations.

Documentation of Methodology

The goal is to provide a reproducible guide for data-driven journalism. This includes clearly documenting datasets (sources, date ranges), explaining the analytical steps and tools used (such as Python, R, and Tableau), and addressing challenges like data gaps and biases.

Word count: 403

Prototype First Steps

For our prototype, we will make an initial analysis of the current housing situation, specifically how the prices have been changing lately. We will also sketch an initial design for our final article, with the topics from each section, and the arrangement of the graphs and

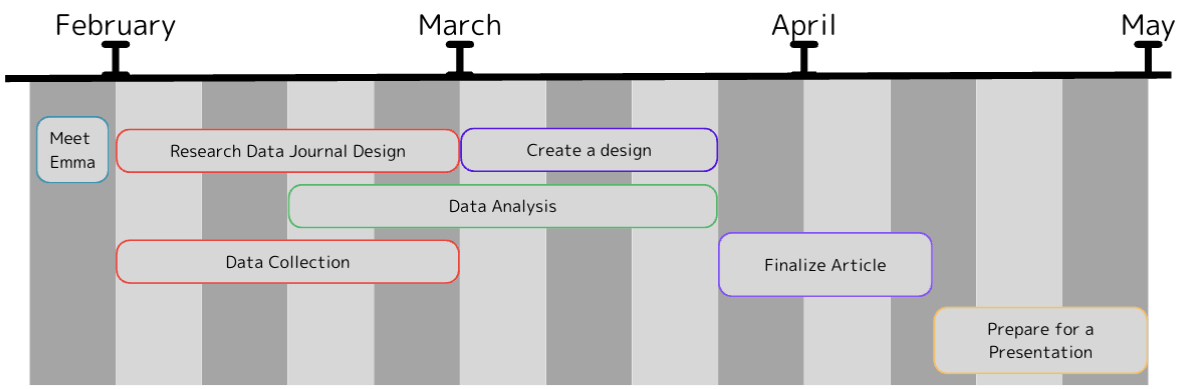
images. This is important because making an entire data analysis from scratch requires applying different contents that we have been seeing throughout the last year in our Formal Analysis classes. By making a simple but effective prototype, we can showcase our intended final product to our #audience and gather feedback on how to iteratively improve it before moving on to the next step—crafting the article with the results from the data analysis.

Word count: 118

Updated Teamwork Direction

Last semester, we collaborated very well by contacting each other through Instagram, as stated in the Team Agreement. However, our meeting was very spontaneous and less scheduled constant meeting. This is due to considering the workload we had for our classes, as it was hard for most of us to balance work and this project. However, we made sure that everything was on track and moved forward. The team survey helped each of us to reflect on our own actions, and we will stick to this team agreement for this semester, hoping we will have more ability to do better.

The one positive aspect of our group collaboration is how motivated we are for this project. When one of us proposes a meeting, we are quick to schedule one and get together. Even if one of us can not meet, we will make sure to update each other so no one falls behind. Two ways in which we can improve our collaboration are first, balance the workload with assignments, projects, and life. What happened last semester was that many people were overwhelmed and the progress of the project got slower as it got towards the end. This is crucial for us to come up with more creative and innovative ideas. The second is for all of us to keep up the motivation. This will largely influence the reply rate and the progress too.



We decided to create multiple roles: team data analysis, team design, and team writing. We will be working together depending on how heavy the workload is, so the allocation of the work will be very flexible. The schedule is created with units of the week, considering if anything happens. Team data analysis consists of Van, Marcel, and Hammon. Design and Writing research is done by Tiago and Yua. We will check on each other with the progress, while Van is in charge of tracking it precisely, like how we did last semester.[1]

Word count: 325

Relevant HCs for the Spring Semester

The most important HCs in developing a data journalism article are: #dataviz, #correlation, #induction, and #organization.

We aim to create visualizations that inform about the data we analyze; the #dataviz framework will help us follow the best practices to translate datasets into understandable images.

Our analysis identifies the correlation between the AI Boom and the Housing Crisis,

[1] #responsibility: It is important for all of us to follow the schedule and to do list by the deadline we set for the success of CCP project. This HC application not only how particularities of responsibility for different members of the group combined to lead to a specific outcome in the project but also details some of the strategies used to target specific facets of responsibility, including learning from the last semester.

therefore, #correlation will guide us through it. Additionally, drawing our conclusions from the data, we will create inductive arguments, and we need them to be strong and reliable enough, so the article is credible (#induction).

Finally, writing a journalist article will require us to apply #organization thoroughly, in order to make our communication clear and informative.

Word count: 113

Updated HSR Categorization & Ethical Considerations

Our project is not considered HSR since we are not having any direct interaction with people. We will not receive any kind of feedback or information from experts, citizens, or testers.

Since there is no direct interaction with people, there aren't many considerations to be made. However, our project aims to provide the readers of the article with a greater understanding of the housing situation in San Francisco at the moment, which can have some considerations.

Taking Utilitarianism into account, for example, we can say that by knowing the situation around them, the citizens can be inspired to help their own community or even to be more gentle with the homeless people they find outside, generating a positive impact in society. Since there is no utility cost in this situation, the overall balance would be positive. Therefore, our project can be considered highly ethical.[2]

Word count: 144

[2] #ethicalconsiderations: In evaluating the ethical implications of our project, we considered how the information we present could shape public perception and decision-making regarding housing in San Francisco. While our research does not involve human subjects directly, we applied a utilitarian framework to assess the broader societal impact, particularly how increased awareness might influence policy discussions or individual attitudes toward housing equity. By framing our work within ethical theory, we ensured that our project aligns with values of transparency and social responsibility, reinforcing the ethical dimensions of data-driven journalism

AI Statement

We used AI to generate a summary of the bullet point we took notes of, in Section 3, which talked about the latest project direction. Other than that, we did not use any kind of AI to protect our learning experience.

Bibliography

We didn't use any resources for the submission of this assignment.[3]

[3] #professionalism: We reviewed multiple times to detect mistakes that could be missed in one's self-proofreading, reviewing all formatting and design choices that might have been changed throughout the writing process, and adhering to the field's professional standards contributed to the overall assessment of the proposal. This allowed us to incorporate our professional habits.