SP-2 Data Mining AI

SOFTWARE REQUIREMENTS SPECIFICATION (SRS)

CS 4850 - Section 03 – Spring 2023

Professor Sharon Perry

02/24/2023

Red Team

Prepared by:

Peter McFarlane, Dev Malani, Manav Desai, Duc Ho

Table of Contents

1 - Introduction

1.1 Overview

Our project is to create a web application that serves as a data visualization tool for the National Crime Victimization Survey (NCVS). The NCVS is a comprehensive survey of crime victimization, with data that spans from 1992 to 2021, obtained from a sample of nearly a quarter million individuals in over 150,000 households nationwide. It includes crimes that are not reported to police, allowing for a more thorough representation of crime victimization in the United States than other crime surveys.

Our web app will mine different variables from the survey and plot their prevalence against various demographics of people. These variables include property crime (theft, burglary, etc.) as well as personal crime (aggravated and sexual assault, robbery, etc.)

We will use various machine learning algorithms such as linear regression and clustering to create detailed analyses of the data.

1.2 Project Goals

Our aim is to be able to identify crime trends in the United States, including 1) which groups of people are most likely to be victims of crime, 2) what kinds of profiles we can establish about the perpetrators of crime, 3) how the prevalence of different types of crime has changed over the last few decades, and 4) which crimes are less likely to be reported to police and what clues the data can provide as to why that is.

1.3 Definitions and Acronyms

NCVS - National Crime Victimization Survey

BJS - Bureau of Justice Statistics

1.4 Assumptions

It is assumed that the data found within the NCVS contains, for all practical intents and purposes, correct information except where explicitly stated within the dataset (survey errors, out of universe, etc.)

It is also assumed that the end user is using a machine that can load web apps.

2 - Design Constraints

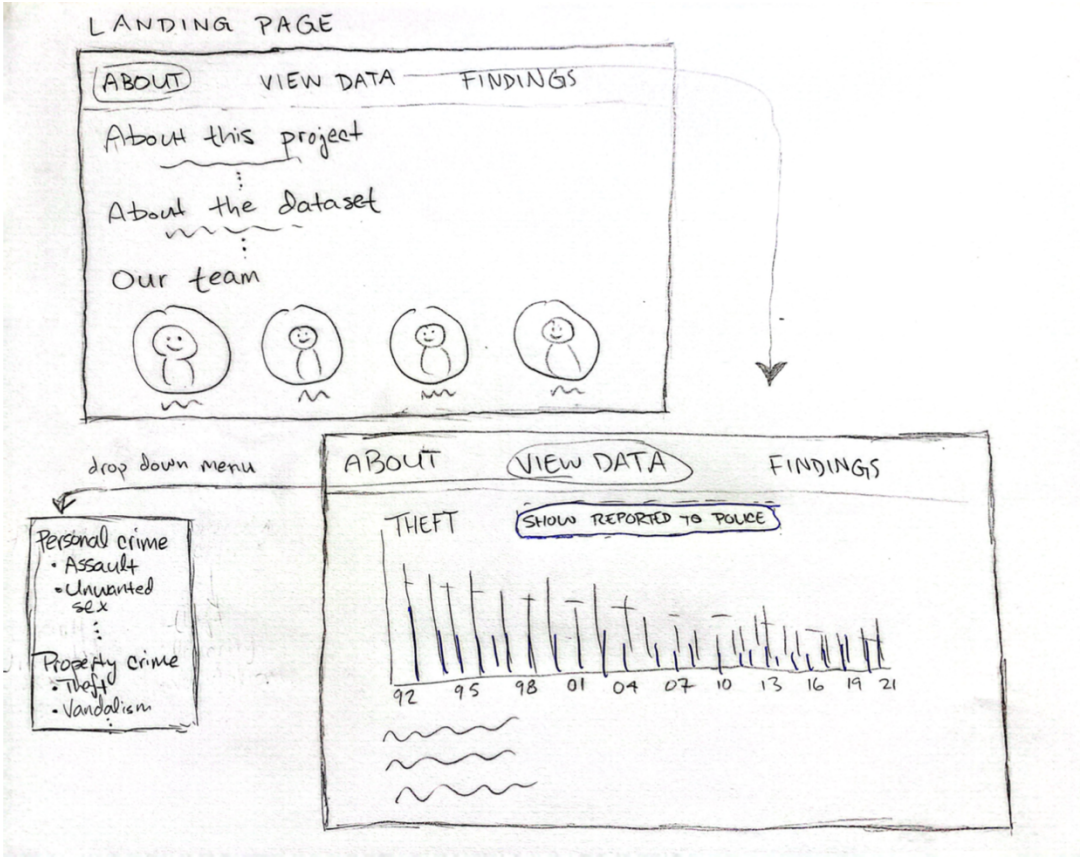
2.1 Environment

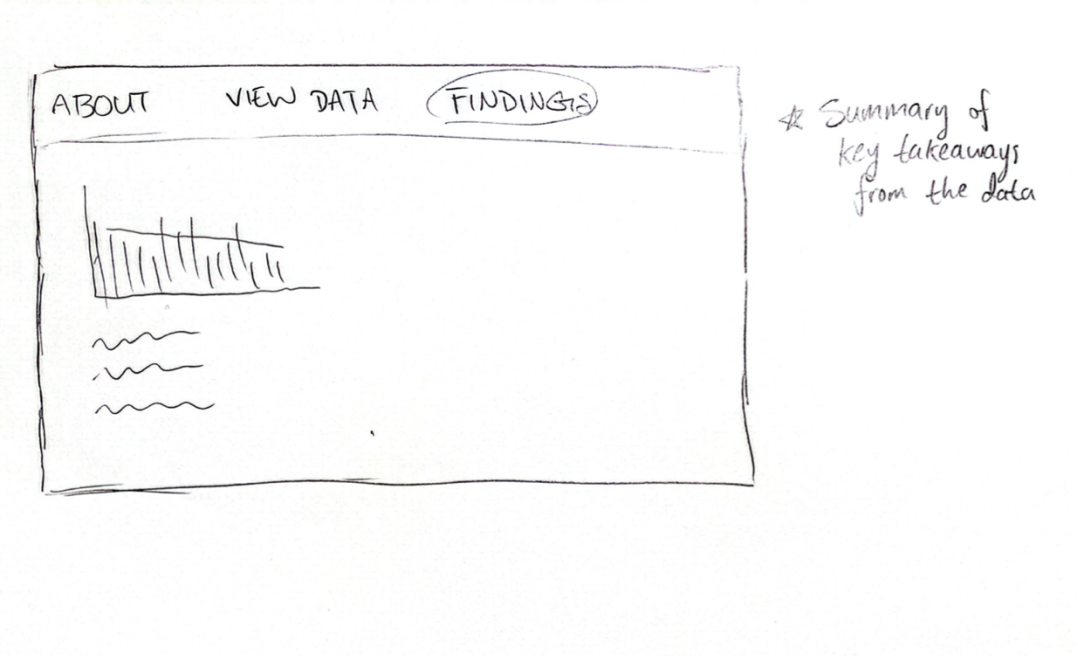
We will be using Python and various Python libraries such as Pandas, NumPy, and Matplotlib to mine the data and generate our graphs. We will use the Django and React framework to build the web app. Our version control will be done through GitHub.

2.2 User Characteristics

The target users for our app are statisticians or social scientists looking for a simple way to visualize the NCVS or otherwise curious individuals looking to learn more about crime in the United States. Perhaps the user is looking to visualize certain data, or perhaps they are simply browsing the data; we will seek to make our app easy to use for both use cases.

3 - Functional Requirements





- About (Landing Page)
    - General information about the project and its methodology.
    - Information about the dataset.
    - Our team and headshots.
    - Contact button at the bottom of the page for users to email us with feedback.
- View Data
    - The core functionality of the application.
    - Allows the user to generate graphs of selected data within a selected time frame.
    - “Show reported to police” overlays the subsection of selected individuals who reported their crime to the police.
- Findings
    - Summary of interesting findings and key takeaways from the data.

4 - Non-Functional Requirements

4.1 Capacity

While the time it takes for operations to complete may vary based on the amount of data being plotted in each operation, we expect that it should take no longer than 5 seconds to generate a graph. The site should be able to handle over 30 concurrent users at this 5 second threshold.

4.2 Usability

Users will be able to provide feedback to the developers via an email survey. We will consider 90% satisfaction or greater an acceptable satisfaction rate.

5 - External Interface Requirements

5.1 User Interface Requirements

The UI will be clean, simple, and professional. It will allow users to select variables via a drop-down menu and generate graphs with the click of a button. We anticipate that north of 90% of users will be able to easily use and navigate the application with fewer than 10 seconds of prior training.

5.2 Software Interface Requirements

Our Webpage is within the Django Project, which will include HTML, CSS to style the webpage, A Web server provided by Django to runs the webpage and to keep our database. We will use the REST API to allow apps to communicate with each other.

5.3 Communication Interface Requirements

The communication interface for our website will be the REST API, which allows apps to communicate over the internet. REST API uses HTTP methods to retrieve, update, and delete data from a server. We can use this interface to access and manipulate data from external systems.

APPENDICES

Pandas Documentation: https://pandas.pydata.org/docs/user_guide/index.html

NumPy Documentation: https://numpy.org/doc/stable/reference/index.html

Django Documentation: https://docs.djangoproject.com/en/4.1/