PW Communications

HomelessR Package

Becca Ebersole, Adam Romriell,
Justin Schultz, Hunter Rogers

Members of this Project

Hunter Rogers - Data Science Major at BYU-I and incoming intern at The Centr. He is graduating on July 20th 2022.

hunterwr1997@gmail.com

Adam Romriell - Data Science Major at BYU-I. He is graduating on July 20th 2022, speaks Russian and is looking for a stable career to help him provide for his family.

adamromriell@gmail.com

Justin Schultz - Data Science Major at BYU-I, graduating July 2022. He is currently seeking a position where we could use his skills in data and analytics.

juschultz33@gmail.com

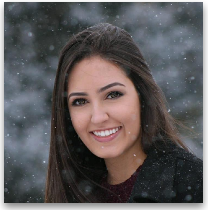
Becca Ebersole - Technical Project Manager of this project. She is graduating on July 20th 2022 with a Bachelors in Data Science. She is moving to Portland, Oregon.

beccawebersole@gmail.com

Faculty Mentors and RBDC Support

Ellie Anderson is the Data Science Liaison at the RBDC. She graduated in 2022 from BYU-Idaho with a degree in Biostatistics and will be leaving for graduate school in August.

- e.anderson@rbdcenter.org



Laura Pires is the Data Science Liaison at the RBDC. She is studying Industrial-Organizational Psychology and Statistics, and will be taking over for Ellie in the Fall.

- l.pires@rbdcenter.org

Ryon Brewer is the Director of Business Development at the RBDC. Before working at the RBDC, Ryon has been a Senior Level Software Executive and CEO of a Software Implementation Company.

-r.brewer@rbdcenter.org

J. Hathaway manages the Data Science Program at BYU-I and is a faculty mentor for these RBDC projects. He has been with BYU-I since 2015 after a 10 year career as a statistician and data scientist at PNNL.

- hathawayj@byui.edu

The goal

- Make homelessness data analysis easy for non-data scientists.
- Expand the possibilities for analysis by integrating external data.
- Future proof data collection with functions that directly access the source.

Data sets

- Funding data
- HUD data
- Census Data
- Crime Data
- Bureau of Economic Analysis (BEA)
- Unemployment Data

Department of Housing and Urban Development (HUD) Dataset

The HUD datasets main purpose is to have the overall homelessness in a given State. Other information that this dataset gathers is on homelessness shelters.

Data preparation

- Read Excel data from a URL
- Remove noise from column names
- Change data types to numeric
- Change state abbreviations to names

Source: https://www.hudexchange.info/resource/3031/pit-and-hic-data-since-2007/

Bureau of Economic Analysis (BEA) Data

This data provides employment and GDP information

Data preparation

- Call data from API
- Convert JSON to dataframe
- Rename columns and remove extra characters
- Pivot data
- Aggregate to state and year level
- Un-abbreviate state names

Source: https://www.bea.gov/about/who-we-are

Source: https://www.icip.iastate.edu/tables/employment/unemployment-states

American Community Survey Data (Census)

Provides insights into population numbers by age, sex, and citizenship status.

Data preparation

- Call data from API
- Pivot the data to long format
- Join column names
- Remove margin of error rows
- Remove noise from column names

Crime Data

The FBI provides the number of crimes in the United States for each state. It breaks down the type of crimes committed and also groups them by violent crime and property crimes.

Data preparation

- Identify the correct link to use per year
- Download csv via link
- Split and fill some row and columns
- Remove noise from column names
- Account for changes in labeling throughout the years

Source: https://ucr.fbi.gov/crime-in-the-u.s

R Package

Static vs Dynamic

Trelliscope/Shiny

We used Shiny to display trelliscopes that we made that the user can look at to answer some of their questions.

We also made it so the user can also create their own trelliscopes without coding.

Showcase of data/Transfer of technology

- The current data 'lives' in the homelessR package. This for ease of access.
- Each data set, except the spending, is derived from a function that goes out and gets the data.
- The BEA and the Census data set functions require an API key to 'query' the API.

What are the next steps?

- Documentation
- Code cleaning
- Improve dashboard