GRAND CANYON
U N I V E R S I T Y™

Scott Maxwell

CST-451 Capstone Project Proposal

Grand Canyon University

Instructor: Professor Michael Landreth

Revision: 1

Date: 07/18/2024

ABSTRACT

CitySurf provides an interface in the form of a web application to view various statistics about different cities in the United States. To create this experience, a full stack application was created to provide a feature rich experience while having a minimal non-overwhelming interface to interact with the data. The goal is to provide a more consumable data visual tools that draw the eye to the conclusion.

Comprehensive graphs and charts are publicly available without a sign in for each city. The graphs representing the various metrics are job industry, salary, employment, health insurance, walk score, home prices, weather, rooms per household, vacancy, and possibly other data points. There will also be community contributed data which will require a login to see add contribute to. These data points include ratings (out of 5 stars) of Pollution, Safety, Friendliness, Attractions. A summary for the city will be made available near the top of the page to help readers get an overall sense of the city.

City Surf will be hosted in the cloud and use high security standards for encrypting the little user data that is collected (username and email) to lower the risk in the event of a data breach. The data itself will be collected via publicly available data sets from the Census Bureau, NOAA, and potentially other reliable sources. The summary for each City will be generated through prompting ChatGPT.

History and Signoff Sheet

Change Record

Date	Author	Revision Notes
07/18/2024	Scott Maxwell	Initial draft for review/discussion

Overall Instructor Feedback/Comments

Overall Instructor Feedback/Comments

Integrated Instructor Feedback into Project Documentation

☐ Yes ☐ No

Project Approval

☐ Professor Michael Landreth

TABLE OF CONTENTS

Project Overview and Project Objectives

The Problem and Background

The few existing sites for browsing city data are not user friendly or consumable. For example, city-data.com offers data in a way that is not interactive nor interesting. There's too much data too consume and it's not easy to get a sense of the city's vibe within a couple of minutes. CitySurf aims to achieve an easier way to look up cities for those who might be looking to move locations or even generally curious as to what city they might be passing through. CitySurf will also bring knowledgeable insights to those seeking to learn a little bit more about their hometown. Another key improvement is to compare cities directly, displaying the data for each city side by side or even in the same graph to show what makes each city unique.

Christian Worldview

CitySurf aims to help individuals find a new potential home. Finding a new home can be frustrating and tedious. From the Christian perspective, we aim to provide honest statistics for each city and to do so in a manner that is helpful as scripture instructs, "Carry each other's burdens, and in this way, you will fulfill the law of Christ" (Galatians 6:2, New International Version). We also look to do such in a way that is enjoyable for the user to make the experience. Paul writes "And do not forget to do good and to share with others, for with such sacrifices God is pleased" (Hebrews 13:15, New International Version). This project will also prioritize the transparency of any user data that is collected and make it a goal to not increase what is collected. In doing this we are loving our neighbor as the greatest command instructs in scripture (Matthew 22:37-39, New International Version).

Project Objectives

- Collect a meaningful number of metrics for the majority of cities in the U.S.
- Create a beautiful interface to display the metrics.
- Create a secure full-stack application that is loosely coupled.
- Allow for user's to sign in to save comparisons they make and contribute to community ratings of cities.

Challenges

- Gathering data from the various sources and mangling the data to account for missing metrics

- Using secure methods for authentication

- Creating security rules for returning responses for web requests

- Creating meaningful graphs and visuals to display the data

Benefits and Opportunities

Users will benefit from the ability to search cities and compare them with each other using an interface that makes it much easier than existing solutions that only display data for each city at a time. This can make re-location easier as users can determine if the city's attributes are right for their needs.

Project Scope

The scope of this project is to provide an interface to display metrics for most of the cities in United States. This interface will be rich of features that make interacting with the data interesting. Data outside of the United States will not be collected and the ability to generate

Stakeholder Name	Role(s)	Responsibilities
Self	Developer and Product Owner	Designing, Developing, and Creating the Application

Work Breakdown Structure											
ID	Task	Dependencies	Status	Effort Hours	Cost	Start Date	Planned Completion	Estimate to Completion	Actual Completion	Resource
1	Collect Data (Data Engineer)	N/A	incomplete	N/A	N/A	N/A	N/A	N/A		
2	Build Front-end (Front-end engineer)	3	incomplete	N/A	N/A	N/A	N/A	N/A		
3	Build Back-end (Back-end engineer)	N/A	incomplete	N/A	N/A	N/A	N/A	N/A		
4	Deploying App (Back-end engineer)	1,2,3	incomplete	N/A	N/A	N/A	N/A	N/A		

Project Success Measures

The Project will be considered a success under the conditions that data for most of the U.S. Cities is collected. Another indicator of success is the number of users who create an account. Lastly, the number of features implemented out of those that are collected will be another indicator of success.

Project Completion Criteria
1 – The number of cities collected out of the 19,495 cities, towns, and villages in United States.
2 – The number of active users who sign in to contribute to the community ratings for each city.
3 – The number of features or requirements implemented out of the those that are collected

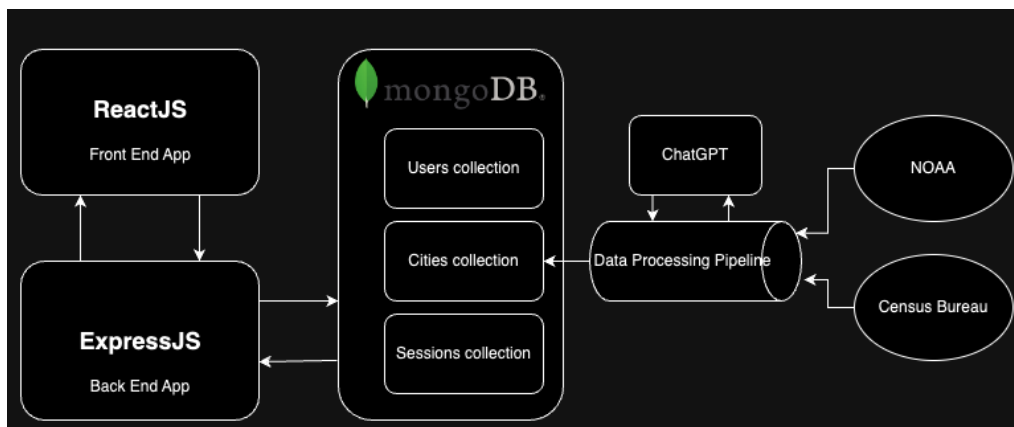| Assumptions and Constraints					
ID	Description	Comments	Type	Status	Date Entered
1	Legal and Licensing	There may be some data that cannot be used due to legal and licensing requirements	Constraint	N/A	7/19/24
2	All required data will be collected on most of the cities in United States	I'm unsure if the data can be collected al within the limited amount of time for this project.	Assumption	N/A	7/19/24

Project High-Level Solution

Introduction:

To provide the large number of metrics for the large number of cities, reliable sources will need to be utilized that are publicly available. A full-stack application will be required for loose coupling. A data pipeline will be needed for the data engineers to process the obtained data and upload it to a database. A backend API will be needed to deliver data from the database to the front-end application. The biggest assumption is that the majority of data will be obtainable via the Census Bureau source, with the second being that it will be manageable to programmatically mangle the data in a desired format for the thousands of cities in the United States.

Solution:

The end user interface will be built using ReactJS. This application will make calls to a back-end application created with ExpressJS. The ExpressJS application will make calls to a MongoDB Atlas database containing Users, Cities, and Sessions collections. The Users collection will be used to store the email and passwords. The Sessions collection will contain user sessions that are currently valid. The Cities collection will make up the majority of the database, containing selected metrics for each city. This data will be processed from the NOAA and Census Bureau with a data pipeline. ChatGPT will be involved to generate summaries for each city with documents being created in the MongoDB collection as the final step.

Project Controls

Risk Management				
Event Risk	**Risk Probability** **(high, medium, low)**	**Risk Impact**	**Risk Mitigation**	**Contingency Plan**
What is the risk?	What is the probability?	What is the impact if the risk occurs?	What can be done to minimize the risk?	What can be done to minimize the impact of the risk?
Data Breach on User Data	Low	User's email and password are leaked.	Implement security for each module of the application.	Encrypt data at rest as much as possible within reason.
DOS Attack	Medium	Service becomes unavailable	Creating security protocols to deny server response to repetitive requests.	Creating redundancy of servers to scale automatically as demand increases

Issues Log								
ID	**Description**	**Project Impact**	**Action Plan/Resolution**	**Owner**	**Importance**	**Date Entered**	**Date to Review**	**Date Resolved**
1	What is the issue?	How will this impact scope, schedule & cost?	How do you intend to deal with this issue?	Who manages this issue?				
2	N/A	N/A	N/A	N/A				
3	N/A	N/A	N/A	N/A				

Change Control Log										
ID	Change Description	Priority	Originator	Date Entered	Date Assigned	Evaluator	Status	Date of Decision	Included in Rev. #
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

End user involvement: End users will be able to submit issues to the public GitHub repository. They can also submit feature requests this way. User involvement will be limited as the stakeholders for this project primarily involve myself. End users who enjoy the site might make suggestions this way but as this a school project there is no initial users to involve until the project has reached a point where it is deployed.

Roles and Responsibilities			
Name	Team	Project Role	Responsibility
Scott Maxwell	N/A	Front-end engineer	Creating the user facing interface that interacts with the back-end
Scott Maxwell	N/A	Back-end engineer	Creating the back-end and setting up the databases
Scott Maxwell	N/A	Data Engineer	Collecting and Organizing Data to Store in the database

Appendix A – References

Bible Gateway. (n.d.). Galatians 6:2. New International Version.
https://www.biblegateway.com/passage/?search=Galatians+6%3A2&version=NIV

Bible Gateway. (n.d.). Hebrews 13:15. New International Version.
https://www.biblegateway.com/passage/?search=Hebrews+13%3A15&version=NIV

Bible Gateway. (n.d.). Matthew 22:37-39. New International Version.
https://www.biblegateway.com/passage/?search=Matthew+22%3A37-39&version=NIV

Appendix B – Copyright Compliance

All MIT licenses are free for any use. CitySurf will make use of these MIT Licensed technologies: NodeJS, ExpressJS, and React. MongoDB Atlas These technologies will make up the full-stack application. ChatGPT will be used to generate summaries for each city. The ChatGPT usage policy permits distributing the output from ChatGPT if it does not harm others. The U.S. Census Bureau Data's terms of service allow the use of the data to be displayed for the purpose of personal or commercial. This data will be used to form most metrics for each city. NOAA's Weather data is public domain and will be used to display the average temperature for each month in each city. Below are the linked Terms of Service and Licensing for each technology or database.

- Node.js License (MIT)
- Express.js License (MIT)
- React License (MIT)
- ChatGPT (Usage Policy)
- MongoDB License (Server-Side Public License)
- U.S. Census Bureau Data Tools and Apps (Terms of Service)
- NOAA's Weather Data (Public Domain)