System Requirements & Design Document

Project Name or Identification

CDI-Viz Dashboard

Prepared by:

Montana Data Consultants

Date:

15 March 2024

Document Versioning Chart			
Version Number	Submit to	Date Submitted	Description
V1	MT-PMI Mentor or Team for editing		
V2	Moodle for grading	27 March 2024	
V3	Sponsor and Team for review & Moodle for revision grading		
V4	Final approval of plan		

Introduction

This document serves as a foundation for the Visualization Dashboard that Montana Data Consultants will be building. With this document, the dashboard's description, requirements, and design will be outlined.

System Description

The system that will be built will tell a "story" about U.S chronic disease indicators from 2001 - 2021 that viewers can "read" at a glance to clearly understand . The system will allow viewers to select the "Disease" and "Indicator" they wish to view. The system will then display a map that is color coded based on prevalence of that disease and indicator within each state. Viewers will be able to quickly comprehend which states have a greater prevalence than others, but will also be able to navigate to individual states and click on it to see more information.

There will be an additional Python program to run new data through a pipeline to clean the data so it is ready to be uploaded into the dashboard.

System Requirements

System Requirement	Description	Priority
Quick Understanding at a glance	Viewers should be able to get an understanding of the data from a quick view	NEED
Differentiate the prevalence by state/year	Viewers should be able to differentiate the prevalence by state/year	NEED
Differentiate the prevalence by other stratifications such as sex and race	Viewers should be able to differentiate the prevalence by sex/race	NEED

User Stories

As a/an (Who)	I want to (What/goal)	So that (Why/Reason)	Notes
Morton Analytics	Display prevalence of disease indicator by state and other Stratifications	Provide example of a potential product that could be created for clients	

Potential Morton Analytics Clients	Display prevalence of disease indicator by state and other Stratifications	Understand what kind of products are available	
General Public	Display prevalence of disease indicator by state and other Stratifications	Gain a better understanding of indicators of chronic disease in the US	

Research Results

Focus groups will consist of Jackson and Ryan from Morton Analytics, Shawn Clouse, and Jason Triche. Each focus group will be asked questions from a predetermined script to gain feedback on the overall UI/UX of the CDI Dashboard.

Research on Tableau will be done throughout the building of the dashboard by the Visualization Specialist in order to gain knowledge and experience within this software.

New information or unforeseen challenges might necessitate revisiting criteria, weighting, or even reopening the evaluation for alternative options; Montana Data Consultants will remain adaptable throughout the evaluation process.

Overview of Systems Used:

System	Strengths	Weaknesses	Reason for Use
Python	- most commonly used language in data science - large community and online resources - open source	- steep learning curve - less statistically oriented than other languages	- preferred language among team members - preferred language of Morton Analytics
Tableau	- most comprehensive data visualization tool - Excellent visuals - free student version	- not very straightforward	- good integration with a variety of platforms - Sponsors familiar with platform

Logical Design

The CDI Data Visualization Tool will be designed so users can view the prevalence of certain chronic disease indicators (asthma, cancer, diabetes, etc.) throughout the United States based on the indicator question, year, data type, location, race, sex, and age. The main visual will be a map of the United States, with each state a different color (using a sunset-themed color scheme) corresponding to the prevalence of the selected indicator and other filters in that state.

Usability

In the context of this project, usability refers to the ease with which users can interact with and derive insights from the data visualization interface. The interface will be intuitive and easy to navigate, allowing users to quickly understand how to interact with the map and access relevant information. The visualization will be responsive and load quickly. The information presented will be clear and easy to interpret. Overall, a high level of usability will ensure users can explore and gain a better understanding of the presence of chronic diseases in the United States.

Navigational Structure and Site Map

The primary interface will be a map of the United States where users can interactively explore the prevalence of various chronic diseases. The interface will allow users to hover over different states to gain more detailed information about the prevalence of the selected disease. Each state in the map will be color coded to represent the prevalence of a specific chronic disease. The prevalence data will be represented using a gradient color scale where darker shades indicate higher prevalence and lighter shades indicate lower. Users will be able to filter data based on: the indicator, year, location, race, sex, and age.

Once the user has picked an indicator, there will be multiple drop down menus on the left side of the screen so the user can filter the data based on the characteristics listed above. The user can look at a specific state, or the user can click on the state on the map itself. Doing so will provide state-level information on the prevalence or rate (depending on the indicator) of the specific indicator the user is looking at. The user will have to specify an indicator to receive any information.

The map itself will be color coded into five classes of prevalence of the chronic disease indicator. The five classes will be color coded in a sunset-based theme. The colors will be yellow, orange, red, pink, and purple, starting at yellow for the lightest prevalence/rate and ending with a dark purple for the highest prevalence/rate. If the data for a specific state is unavailable, it will be coded in gray.

Another drop down menu on the left side of the screen will allow the user to select which year of data they are looking at. The user will also be able to choose to filter by specific age ranges (18-44, 45-64, >=65), sex (male or female), and race/ethnicity (white non-hispanic, black non-hispanic, Hispanic, asian non-hispanic, Hawaiian or Pacific Islander non-Hispanic, American Indian or Alaska Native non-Hispanic, and Multiracial non-Hispanic).

There will be another drop down menu title “Generate figure” and the user can choose to generate a table or graph based on their chosen filters.

Proposed Home Page and/or User Interface (UI) layout for the Information System

The proposed home page will feature a blank map of the United States with a drop down menu above the map, titled “Please choose a chronic disease indicator to view”. The first drop down menu will allow the user to pick which chronic disease indicator the user would like to filter by.

(Please note that the wireframe mock-up is a basic plan, details will be added upon creation of the dashboard).

CDI-Viz Dashboard

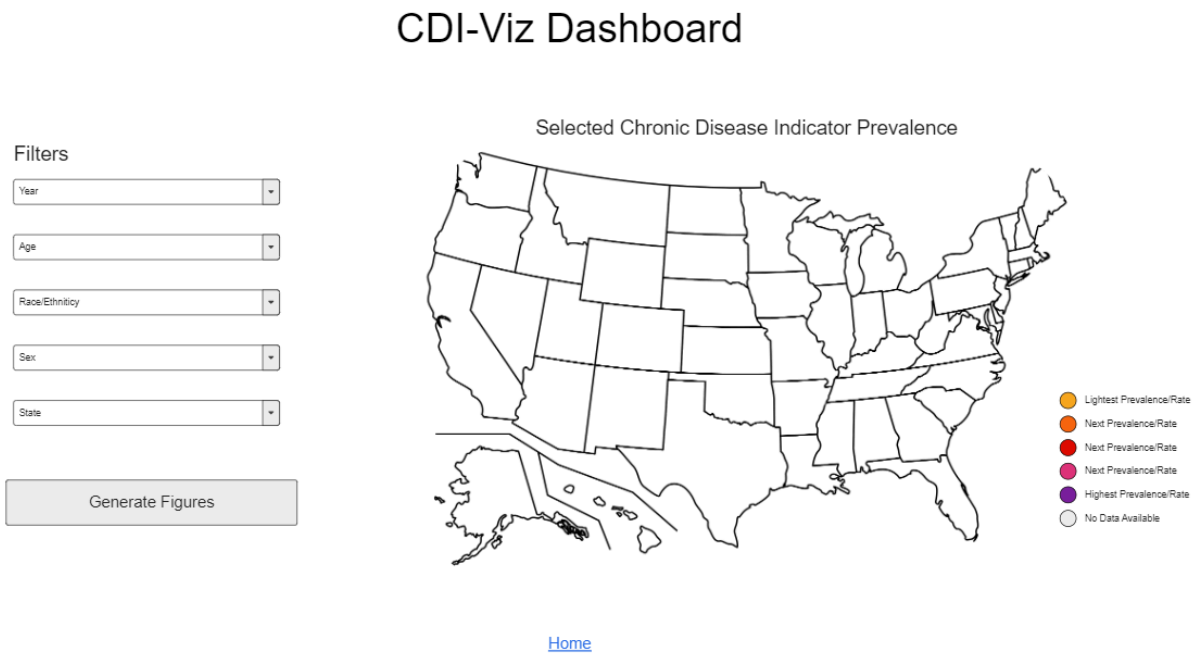
About the Dashboard

Please choose a chronic disease indicator to continue:

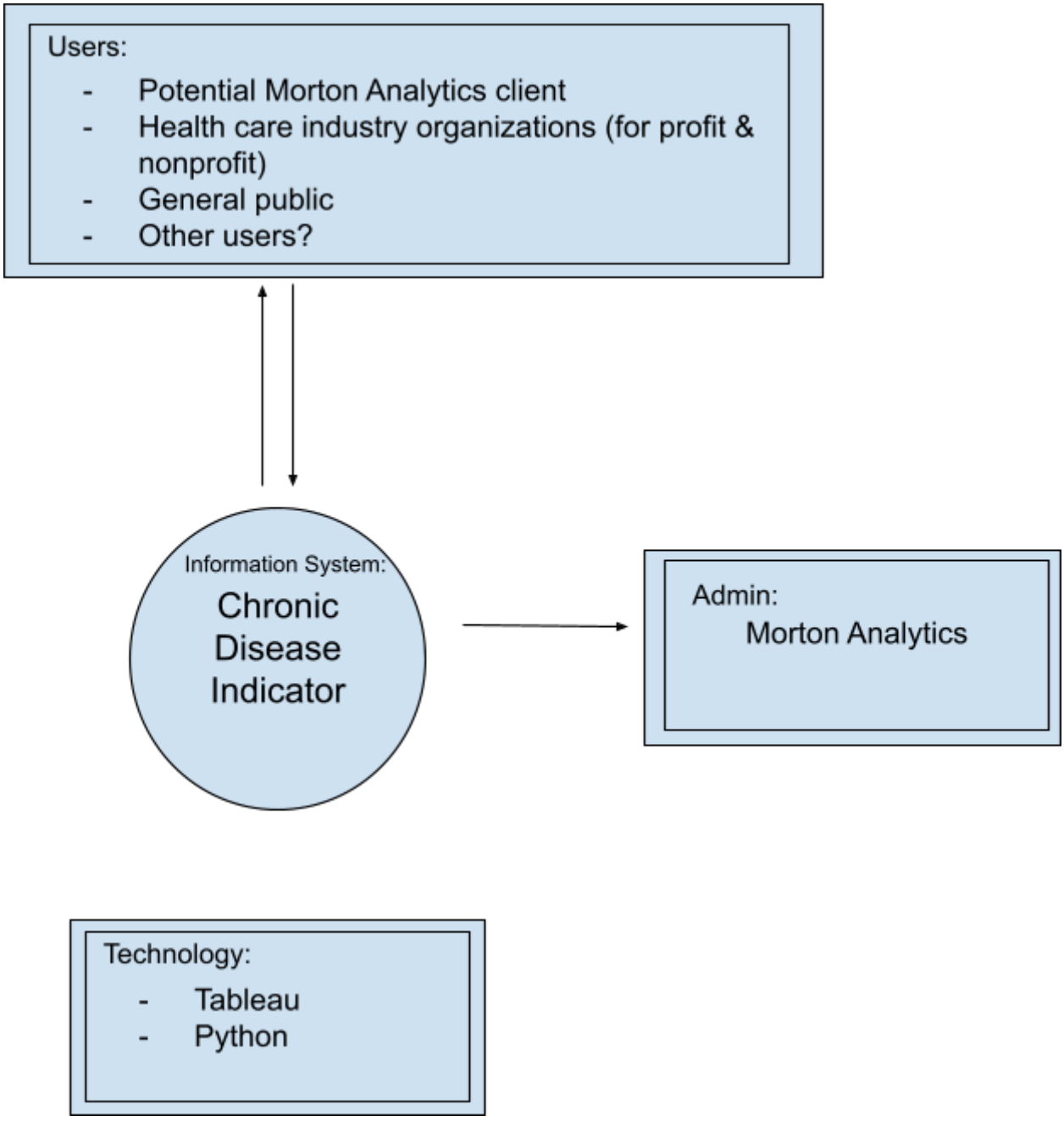
Select

Design Diagrams

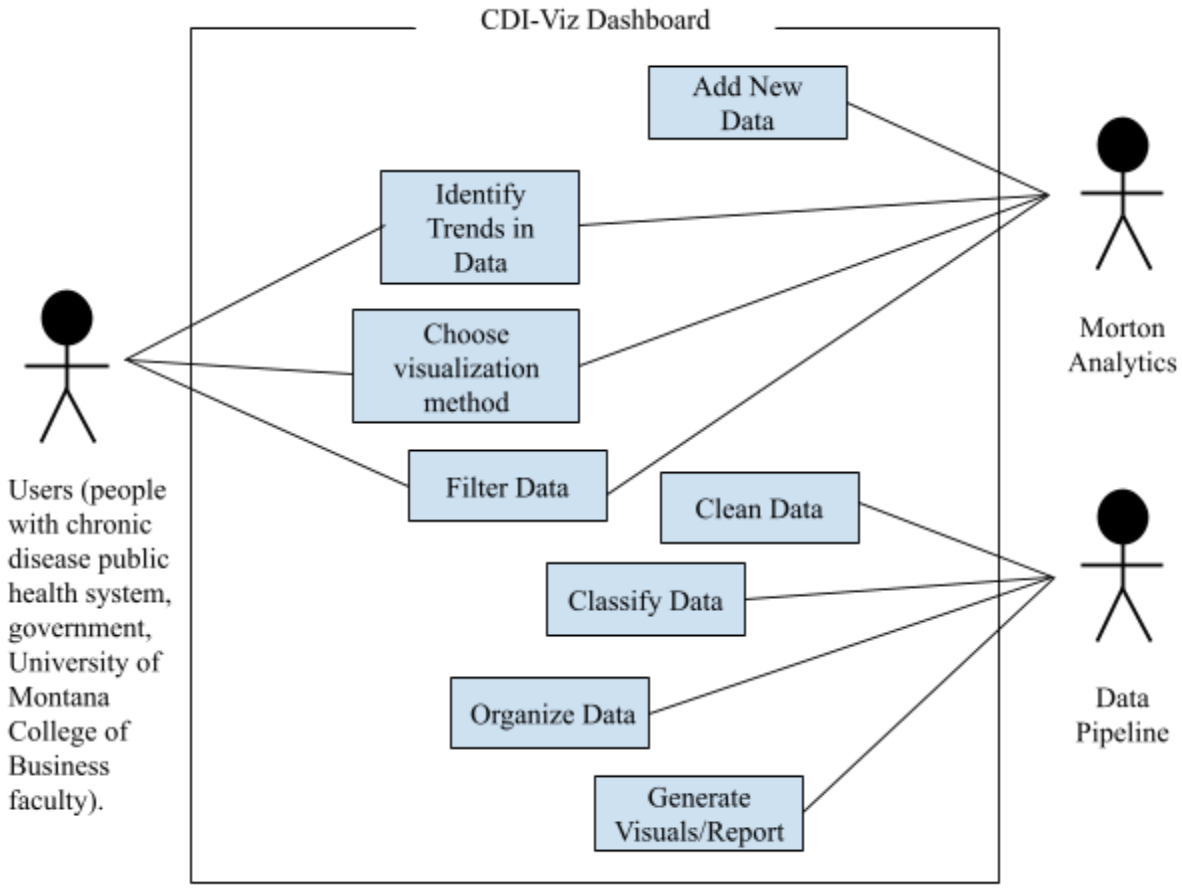
Once the user selects what disease indicator they wish to look at, they will then be able to select additional filters. There will also be an option to go “Home” so they may return to the beginning.



Context Level Data Flow Diagram



Use Case Diagram



Physical Design/Data Architecture

The data architecture of this project aims to provide a robust foundation for visualizing the prevalence of chronic diseases across the United States, incorporating demographic factors such as race, sex, and age. This section of the systems document outlines the design, components, and considerations involved in structuring and managing the data required for the visualization.

Server Used

The servers used in creating this visualization are Python and Tableau. These servers were chosen based on the team's experience with these servers and the system requirements provided

by Morton Analytics. Python will be used to create the data pipeline, and clean data. Tableau will be used to create the visual dashboard utilizing the cleaned data set.

Database

The database for this visualization is a single public dataset available at U.S. Chronic Disease Indicators | Data | Centers for Disease Control and Prevention (cdc.gov). The data comes as a csv file containing 1185676 rows and 34 columns.

Column Name	Description	Type
YearStart	Starting Year	Number
YearEnd	Ending Year	Number
LocationAbbr	Location abbreviation	Text
LocationDesc	Location description	Text
DataSource	Data source abbreviation	Text
Topic	Chronic disease Topic	Text
Question	Question full-length text	Text
Response	Response	Text
DataValueUnit	The unit, such as $, %, years, etc.	Text
DataValueType	Data type such as prevalence or mean	Text
DataValue	Data value such as 12.4, or Category 1	Number/Text
DataValueAlt	Equal to Data Value but formatting is numeric	Number
DataValueFootnoteSymbol	Footnote symbol	Text
LowConfidenceLimit	Low confidence limit	Number
HighConfidenceLimit	High confidence limit	Number
StratificationCategory1	The category of the	Text

	stratification, such as Gender	
Stratification1	The stratification within the category such as Male or Female	Text
StratificationCategory2	The category of the stratification, such as Gender	Text
Stratification2	The stratification within the category such as Male or Female	Text
StratificationCategory3	The category of the stratification, such as Gender	Text
Stratification3	The stratification within the category such as Male or Female	Texy
Geolocation	Geolocation code for mapping	Point
LocationID	Location identifier	Text
TopicID	Topic identifer	Text
QuestionID	Question identifier	Text
ResponseID	Identifier for the response	Text
DataValueTypeID	Identifier for the data value type	Text
StratificationCategoryID1	Identifier for stratification category 1	Text
StratificationID1	Identifier for stratification 1	Text
StratificationCategoryID2	Identifier for stratification category 2	Text
StratificationID2	Identifier for stratification 2	Text
StratificationCate3goryID	Identifier for stratification category 3	Text
StratificationID3	Identifier for stratification 3	Text