|  |
| --- |
| PROJECT PROPOSAL - Group PROJECT 2  Visualization of Data Scientist Opportunities |
| CHRIS GILBERTJAMES DODSONPETER HOLDER PRAGATHI MADUSHINIE |



**Visualization of Data Scientist Opportunities**

**Project Proposal for Group Project 2**

INSTRUCTIONS

Summary page:

a. Project title: Visualization of Data Scientist Opportunities

b. Name of implementing person(s): Team Purple

c. Project location: Data Analytics Boot Camp - WashU

d. Proposed starting date: 08/11/2019

e. Project duration: 2 weeks

PROJECT REQUIREMETS:

Project Description:

* Your task is to **tell a story** through data visualizations.
* Focus on providing users an **interactive means** to explore data themselves.
* Prepare a **10-minute presentation** that lays out your theme, coding approach, data
* munging techniques, and final visualization.
* You may choose a project of any theme, but we encourage you to **think broadly.**
* You will have **ample time in class** to work with your group, but expect to put in **hours outside of class** as well.

Specific Requirements of the Project:

* Your visualization must include a Python Flask powered RESTful API, HTML/CSS,

JavaScript, and at least one database (SQL, MongoDB, SQLite, etc.).

* Your project should fall into one of the below four tracks:
  + A custom “creative” D3.js project (i.e., a nonstandard graph or chart)
  + A combination of web scraping and Leaflet or Plotly
  + A dashboard page with multiple charts that update from the same data
  + A “thick” server that performs multiple manipulations on data in a database prior to visualization (must be approved)
* Your project should include at least one JS library that we did not cover.
* Your project must be powered by a data set with at least 100 records.
* Your project must include some level of user driven interaction (e.g., menus,
* dropdowns, textboxes).
* Your final visualization should ideally include at least three views.

The Project Proposal for the Group Project a has three sections:

I. What Story are we trying to tell from the data?

II. What tools are we using to implement the story we are telling?

III. How we implement the process?

I. WHAT STORY ARE WE TRYING TO TELL FROM THE DATA?

Our project will be a user-based portal to visualize job listing for data scientists across the United States. The focus will not be only on the jobs available, but a mapping of skill sets required or requested for the positions in our dataset. The visualizations may include:

* Heat map of positions available
* Top required/requested skills in certain geographies
* Matching of skill sets the user has with available openings

The intent of this project is to be a useful tool to assist not only in job searches, but also in determining what skills are in high demand. This tool will also be able to determine if there are certain skills that are in higher demand in certain location.

II. WHAT TOOLS ARE WE USING TO IMPLEMENT THE STORY WE ARE TELLING?

Front End Tools: HTML, CSS, BootStrap, JavaScript

Data Visualization Tools: JavaScript – D3, leaflet

Database and server: PostgreSQL, Python, SQLAlchemy, Flask

Our project will focus on these areas:

○ A custom “creative” D3.js project (i.e., a nonstandard graph or chart)

○ A combination of web scraping and Leaflet or Plotly

○ The extra JS library we are using is undetermined at this time

○ Our data set will include more than 5,000 job listings and skills

○ User driven interactions which may include menus, dropdowns, and textboxes

III. HOW WE IMPLEMENT THE PROCESS?

This project will be divided into specific tasks related to the functionality of the project. Each team member will be responsible for specific components of the project. We will determine the tasks and assign by the end of class on August 15th.