Data228

Big Cloud

Yixin Huang
Chloe Ngo
Minh-Tam Pham
SeungMin Yoo

Project Plan

Version 1.0

Revision History

Date	Version	Description	Author
05/10/2021	Final	Front end	Yixin Huang
05/14/2021	Final	Analysis	Minh-TamPham Chloe Ngo
05/16/2021	Final	Predict	SeungMin Yoo
05/16/2021	Final	ETL	Yixin Huang

Table of Contents

Introduction

1.1 Purpose of this document

The purpose of this document is to provide a detailed overview of the steps taken during the development of the Covid-19 Labour Force analysis application, by and for Team Big cloud.

1.2 Intended Audience

This document shall be used in all phases of the project as a guideline. Intended audiences of this project are all project stakeholders:

- project supervisor
- project leader
- team members

1.3 Scope

This document defines the project plan of the Covid Job rates analysis application. The overview includes objectives of the project, organization of the project team, development process that is going to be used during the project, assessment of possible risks, communication used between project stakeholders and project plan that includes time schedule and activity plan.

1.4 Definitions

1.4.1 Definitions

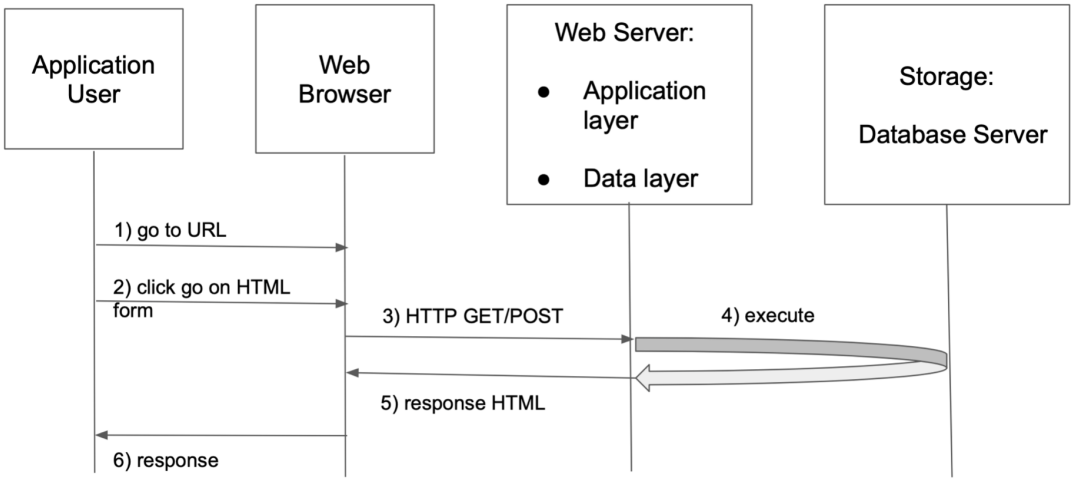
Keyword	Definitions
<Name>	The name of the project
Yixin Huang	A person in charge of supervising the project
Chloe Ngo	A person in charge of organizing the team and communicating with the project supervisor
SeungMin Yoo Minh Pham	An active member of the team responsible for making the job done
Milestone	A time in a project that marks the end of a project phase or the completion of an important deliverable.

Background and Objectives

To say Covid-19 had an impact on the economy is an understatement. The United States faced peak unemployment rates of 14.8 percent in April of 2020, and recovery is certainly on everyone's minds. Our project aims to help researchers, private and public, analyze the recovery of the labor force and discover insights about how the economy is recovering. We want to zero in on the labor market's efforts to meet pre-Covid norms, and we will be using a dataset that analyzes the rate of new job postings since March 2020.

The rates are relative to pre-covid times (set at 1) and will measure if there are less (rate <1) or more (>1) job listings compared to normal, pre-Covid times.
Additionally we will ingest an additional source of data measuring Covid mortality [1]

Architecture & High Level Design



Our web application is designed with three parts: front-end, middle-tier server, and back-end database. For the full stack process, when the user types in the login URL and hits 'go', the browser will find and request from the server of that particular page. Then the server responds by extracting files from the back-end database and sending those files over to the browser. So the user can see the login web page at the front-side.

Organization

1.5 Project group

Name	Initials	Responsibility (roles)
Chloe Ngo	CN	Team Leader
Minh Pham	MP	Member
Seung Min Yoo	SY	Member
Yixin Huang	YH	Member

1.6 Customer

Our project aims to help people better understand the labor market's recovery rates by examining the rate of new jobs per day, relative to pre-covid times. The target customers are listed below:

- Government workers
- Economists
- Researchers
- Students

The insights of this project could be helpful in broad research, decision making at the public and private level, and anticipating the future of the US economy.

Development process

- Flask stack: HTML- Python - FLASK- MySQL(PHP ADMIN)
- On the front-side, we pick HTML as the language to communicate through HTTP requests.
- On the server-side, we use the python package FLASK, Jinja2 to run the server and respond to HTTP requests.
- The back-end databases, we chose MySQL PHP myadmin as our database since it's free under a certain limited storage and easy to access to our front-end by using flask.

Deliverables

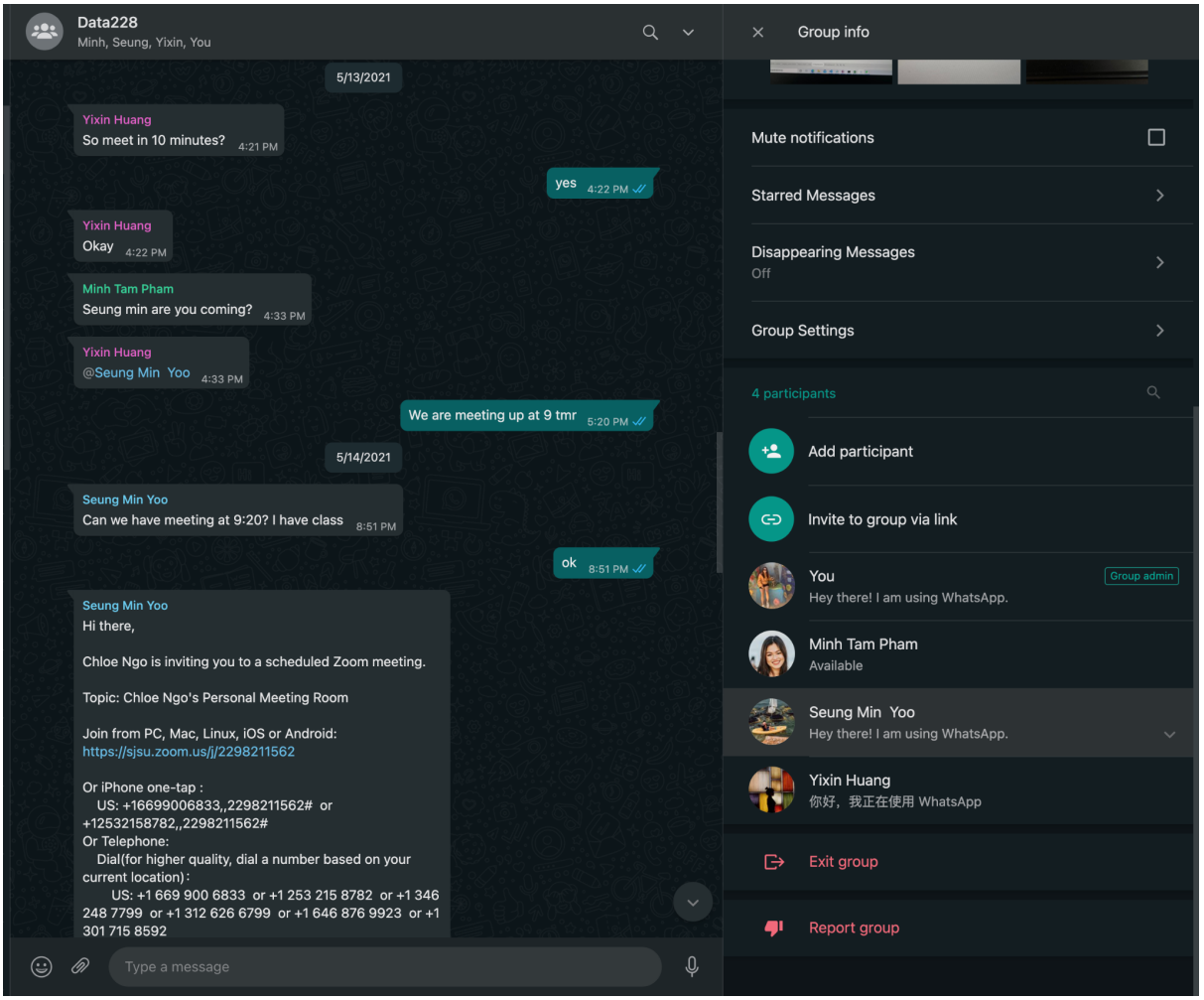
Name	Output	Planned week	Late +/-	Delivered week	Notes
Project Code	**Github**	**05/05**	**-**	**05/10**	**on time**
Project Presentation	**ppt**	**05/05**	**-**	**05/10**	**on time**
Project Report	**pdf**	**05/23**	**-**	**05/23**	**on time**
Project Proposal	**pdf**	**05/23**	**-**	**05/23**	**on time**

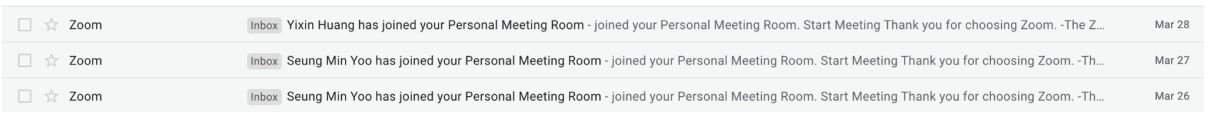
Project risks

Possibility	Risk	Preventive action
High possibility	going over budget	We will build our own warehouse using myPHPclient and build our front end instead of using AWS.
Medium possibility	poor knowledge on app-building	meet with ISA for guidance

Communication

Communication primarily takes place in a group chat, WhatsApp, where members plan out dates/times of zoom meetings and discuss overall plans.
The team met up at least twice a week to discuss the project, collaborate on code together, and problem solve when errors arise.





1.7 Collaboration

The team will split the work evenly, so we all have the opportunity to learn about each portion of the project, but each individual will have a primary responsibility. Seungmin will be responsible for the ETL process, Minh will implement the data analysis, while Chloe and Yixin will oversee the development of the front end of the application.

1.8 Git

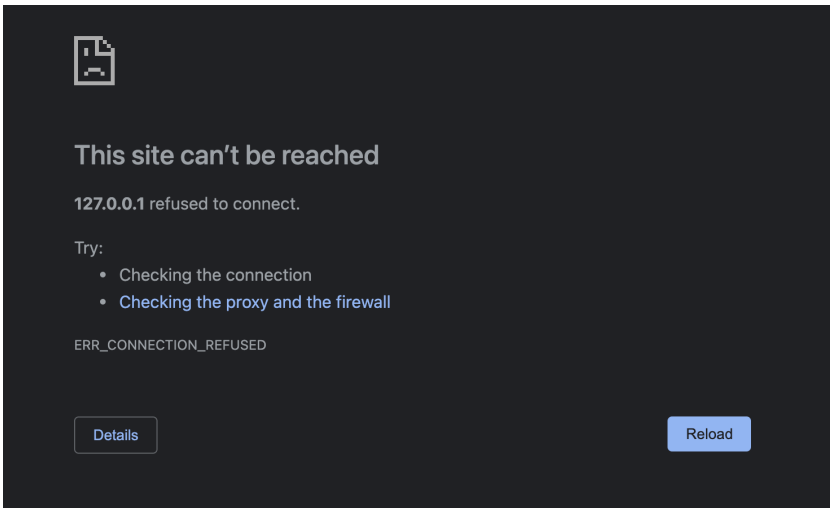
All source code and finished documentation will be uploaded to Github repository.
Repository URL: https://github.com/chloe-chau-ngo/BigCloud

Project plan

1.9 Time schedule

Id	Milestone Description	Responsible Dept./Initials	Finished week (plan)	Finished week (actual)	Notes
1	Find dataset	MP	02/28	03/28	finalized data choice
2	Switched from AWS to independent application building	SY,CN,YH ,MP		04/15	unplanned. AWS too expensive, over 4,000 in charges
3	Build application outline	YH	03/28	03/28	
4	Data cleaning	CN	04/11	04/11	
5	Exploratory analysis	SY	04/11	04/11	
6	Analytics + visualization	CN, YH	04/11	04/18	
7	Add SSO	YH	04/18	04/11	
8	Added new data, ETL	MP, YH	04/25	04/25	
9	added new pages to app	YH	05/2	05/02	
10	Prediction	SY	05/09	05/09	
11	Evaluation	SY, MP, CN, YH	05/09	05/09	Evaluation before submit
12	Commit code to Github	SY, MP, CN, YH	05/10	05/16	final commits

1.10 Test cases

Test No.	001	**Phase:**	1	**Author:**	<username>	Date:
Test Category:		**Launch Website**				
Software Product:						
Test Title:	Testing web app run					
Test Purpose:	To load the data to the database					
Test Setup:	running command on anaconda and command line					
Prerequisites:	we learn that we must initially data					
Procedure:	running command on anaconda and command line					
Checks:	check output					
Expected Results:	open app					
Result:	webpage said error 					
Reason for Failure:	opening app from anaconda, or command line, before opening mySQL database first					
Remarks:	Before running the app, we must test that the database has been linked to the database. Otherwise it will fail					

References

[1] Centraal Bureau voor de Statistiek. “3,9 Duizend Mensen Overleden Aan COVID-19 in December 2020.” *Centraal Bureau Voor De Statistiek*, Centraal Bureau Voor De Statistiek, 6 Apr. 2021, www.cbs.nl/nl-nl/nieuws/2021/14/3-9-duizend-mensen-overleden-aan-covid-19-in-december-2020.