Deliverable 1

1) Team

a) Members – Deeksha Umesh, Pruthvi Kasula, Shalini Marneni, Venkateshwar Reddy Sankyapalli, Gunakar Maddi

b) Communication plan to include project artifact repository - <https://github.com/dumesh09/bigdata>

2) Selection of data to analyze from the Open Data Registry for Amazon Web Services https://registry.opendata.aws/

<https://registry.opendata.aws/us-hiring-rates-pandemic/>

3) Business Problem or Opportunity, Domain Knowledge (link to information on domain relative to data, problem or opportunity)

This dataset provides daily updates on the volume of US job listings filtered by geography industry job family and role; normalized to pre-covid levels.These data files feed the business intelligence visuals at covidjobimpacts.greenwich.hr, a public-facing site hosted by Greenwich.HR and OneModel Inc. Data is derived from online job listings tracked continuously, calculated daily and published nightly. On average data from 70% of all new US jobs are captured, and the dataset currently contains data from 3.3 million hiring organizations.Data for each filter segment is represented as the 7-day average of new job listings for a specific date, expressed as a percentage of the corresponding value on March 1, 2020.

Data Specification link:

<https://greenwichhr-covidjobimpacts.s3.us-east-2.amazonaws.com/ghr_data_specs_covid_public.pdf>

The Greenwich.HR data platform is the largest and fastest source of real-time labor market data. Greenwich.HR clients are using our high-definition data platform to take their workforce analytics efforts to the next level.

4) Research Objectives and Question(s) (what you are trying to describe or predict with the data)

Objective:

Descriptive questions:

* What is the impact on job postings in each state.
* How different industries have adapted to COVID.

Predictive questions:

* overall job listings growth