**Assignment 1**

**Information Seeking**

1. **DATASET 1:**

Kaggle (2017), *Kickstarter Projects (3)* [Data file]. Retrieved September 10, 2018, from <https://www.kaggle.com/kemical/kickstarter-projects>

**LICENSE:**

The dataset is released under CC BY-NC-SA 4.0. The license can be found under the following link:

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Under the license, I am free to:

Share – Copy and redistribute the material in any medium or format.

Adapt – Remix, transform, and build upon the material.

**DESCRIPTION:**

Crowdfunding has become a very important aspect to many up and coming businesses these days. Not many start-ups want to depend on traditional VCs to stay afloat in their industry. Crowdfunding platforms are getting more and more popular and prove to be reliable methods to fund your projects and ideas.

Kickstarter, an American public benefit corporation founded in 2009, has successfully funded 150,240 projects till date,with total dollars pledged standing at a hefty 3,883,983,201 USD.1 Kickstarter is quite the successful crowdfunding platform owing to it being different from other websites; the video-as-part-of-the-campaign-feature.

The dataset contains information of over 300,000 Kickstarter projects and their information like:

1. Name
2. Category
3. Goal
4. Deadline
5. Date Launched
6. Backers
7. USD Pledged

The potential data users for this data would be people who are interested in starting their own company but would require an idea of the competition and past experiences of similar start-ups. Once they get a better picture of what categories get most backers and funding they might wish to pivot their business idea for higher chances of success.

**POTENTIAL QUESTIONS:**

1. Which categories are the most popular for crowdfunding?
2. In what category do we observe the most successful projects?
3. Does the country pledged from make any difference to the status of the project?
4. **DATASET 2:**

State of Maryland (2018), *Choose Maryland: Compare States – Quality of Life* [Data file]. Retrieved September 11, 2018, from <https://healthdata.gov/dataset/choose-maryland-compare-states-quality-life>

**LICENSE:**

This dataset is published under the ODC Open Database License (ODbL). The license can be found under the following link:

<https://opendatacommons.org/licenses/odbl/summary/>

Under this license, I am free to:

Share: Copy, distribute and use the database

Create: Produce works from the database

Adapt: Modify, transform and build upon the database

**DESCRIPTION:**

In today’s fast paced world, the general well-being of individuals and societies is often neglected. Life satisfaction is disrupted by too many factors; not having a balance between work and personal life, compromising on physical health and family, not having complete access to the best education etc.

This dataset contains information on all states of the United States of America like the following:

1. Number of National Parks
2. Number of National Historic Landmarks
3. Percent of Households with Broadband Internet
4. Well Being Index
5. Active Physicians per 10,000 Population

This data can be used by students and researchers involved in psychological studies who may require data on the well-being index of all the states in USA. People involved in the health and human services industry can use this data to analyse whether there is any correlation between the stable income households and the well-being index.

**POTENTIAL QUESTIONS:**

1. Which state in the United States has the most optimum well-being index?
2. Is there any association between the well-being index and the number of physicians in the state?
3. Is there any pattern to the Arts, Entertainment, Sports and Media Employment geographically?
4. **DATASET 3:**

Kaggle (2017), *US Jobs on Monster.com* [Data file]. Retrieved September 11, 2018 from <https://www.kaggle.com/PromptCloudHQ/us-jobs-on-monstercom>

**LICENSE:**

This dataset is published under the CC BY-SA 4.0. The license can be found under the following link:

<https://creativecommons.org/licenses/by-sa/4.0/>

Under this license, I am free to:

Share – copy and redistribute the material in any medium and format.

Adapt – remix, transform, and build upon the material for any purpose, even commercially.

**DESCRIPTION:**

The US Job market is consistently booming in various industries. There are various job posting platforms out there that are continuously being updated with job postings. Companies of all sizes and popularity use such platforms. It is interesting to see how job applications, a procedure that use to depend on face to face interaction solely, has now been digitized like the rest of the World.

The dataset contains information on:

1. Country
2. Job Description
3. Organization
4. Date Added
5. Job Title
6. Job Type
7. Sector
8. Salary
9. Location

The versatility of this data is remarkable since it can be used by the general job-seeking public, people who already have jobs but want to shift and companies who would like to keep a track of the hiring activities of their competition.

**POTENTIAL QUESTIONS:**

1. What sector is the most popular on Monster.com?
2. Is there any pattern between Salary and Job Type? Which companies pay higher wages and for what sector?
3. Is there any pattern emerging geographically with respect to salary and job types?

References:

1. Kickstarter Stats, Kickstarter website, <https://www.kickstarter.com/help/stats>