**Team Red Line – District Data Incubator MVP Description**

Felicia Charlot, Ian Cummins, Wat Hughes, Wayne Liu

**Project Overview**

The District Data Incubator (DDL) Program is a structured 3-month project development program for working professionals in the DC area with an interest in applied data product development.

The objective of Team 1 is to use Bureau of Labor Statistics (BLS) data to create a basic product that can be used by occupational safety professionals to inform the design and delivery of workplace injury/illness prevention programs. The product is intended to extend their ability to design more effective programs by enabling novel and easily accessible visualizations of BLS data on the worker and case characteristics for reported occupational injuries and illnesses. The product will also be designed to enable BLS to highlight specific trends or thematic issues, such as risk factors for older workers.

For this program, we will be working in consultation with BLS. The product, and all related materials, code and interface prototype, will only be shared with our BLS and DDL sponsors; any public releases of the product on the BLS or other websites are beyond the scope of this project.

**Project Scope**

This project is a variant of the BLS’ “Occupational Injuries by Part of Body” project proposal. Similar to the original proposal, our project will be focused on creating a practical visualization tool that professionals can use to inform their prevention efforts. While our proposal will involve case characteristic data, including parts of the body injured, our visualization will not be organized around human body parts. Given the short length of time, developing an integrated recommendation engine for safety equipment is also beyond the scope of this project.

For the Incubator Program, we have been asked to develop a functioning Minimum Viable Product (MVP) by October 24, 2015. An MVP is a basic version of a product with a limited number of functions that are highly valued by a target user group that can be released as early as possible to test viability.

For this project, our target users are occupational safety professionals that design, deliver training and audit workplace safety programs, either internally as employees within a company or externally as consultants or auditors. We chose this user group because they are in a position to directly influence corporate policies and practices; BLS has also confirmed that safety professionals are one of the most active and sophisticated users of their data.

We understand that safety professionals have demonstrated an active interest in both high-level data on the incidence of injuries and illnesses, as well as more detailed data on worker and case characteristics. We also understand that one of their main motivations for accessing BLS data is to compare incidence rates and to monitor outcomes within a specific industry.

We further assume professionals would be most interested in a product that provides access to relevant granular data related to reported injuries, clustered in a meaningful way for specific types of users and made available through easily understood visualizations that enable professionals to test their own assumptions.

Therefore, our MVP will focus on enabling meaningful visualizations by industry of the following variables:

* **Industry (Independent)**
* **Age Range (Worker Characteristic)**
* **Nature of Injury/Illness (Case Characteristic)**
* Days Away From Work (Case Characteristic)
* Gender (Worker Characteristic)
* Event (Case Characteristic)
* Part of Body (Case Characteristic)
* Occupation (Worker Characteristic)

Limited user testing of the visualization prototypes will be conducted to ensure data is presented in a meaningful way.

**Technical Plan**

Our team will use R data analytics software to clean, summarize, analyze and program. The code used to read, organize and analyze the data will be stored on the DDL’s GitHub account, which is publicly available.

We will use Shiny, an R Studio application, to create a web application prototype. We will also enable use of Google Analytics that can be used to collect data on product users, a key Program requirement. The web application interface will only be shared with BLS, DDL and potentially a small number of sample users.