

Meeting Minutes -- IS607 Project 3

Date: October 20, 2016

Time: 9:15 PM

Attendees: Judd Andermann, Ka Man Chan, Marco Siqueira Campos, Scott Ogden, Sharon Morris, Tahla Muhammad, Yefei Li

Discussion

- Data
 - LinkedIn, KDNuggets and Dice.com data have been scraped and cleaned
 - All data and code have been posted to Github
 - Jose will use R-script to pull the data directly into MYSQL from R instead of creating CSV files
 - Jose will pull all 3 data sources into a data table
 - Data will be aggregated by skills categories
 - Missing values will be coded as NA
- Database
 - Three tables
 - Skills table, ID and skill
 - User profile table
 - Jobs tables
- Next meeting
 - Saturday October, 22 -- 6:00PM
 - Agenda
 - Analysis review
- Tasks

Task	Person(s) assigned	Deliverable	Due Date
Setup collaboration tools	Sharon	Github repository	10/16/2016
Identify data sources	Team		10/16/2016
Collect and prepare data for cleaning <ul style="list-style-type: none">● LinkedIn● Twitter● Kdnuggets/DICE	<ul style="list-style-type: none">● Judd● Tahla● Ka Man	<ul style="list-style-type: none">● Code● Data in R	10/18/2016
Tidy data	Judd/ Kaman	<ul style="list-style-type: none">● Code● Data in R table	10/21/2016

		/data frame	
Format data in relational database	Jose	<ul style="list-style-type: none"> • SQL code 	10/21/2016
Statistical analysis & Presentation	Marco/Scott/Bin	<ul style="list-style-type: none"> • Code • PowerPoint 	10/23/2016