EdTech Consulting Report

• • •

MIE1624 - December 2020

Group 14

Overview

1. Re-Designing MIE1624
   a. Data Analysis
      i. Most Commonly Used/Desired Skills
      ii. Skill Clustering and Dependency Modelling
   b. Course Curriculum Design

2. Designing a "Master of Data Science and Artificial Intelligence" at UofT

3. Identifying Data Science Opportunities in Canada Using Review Data
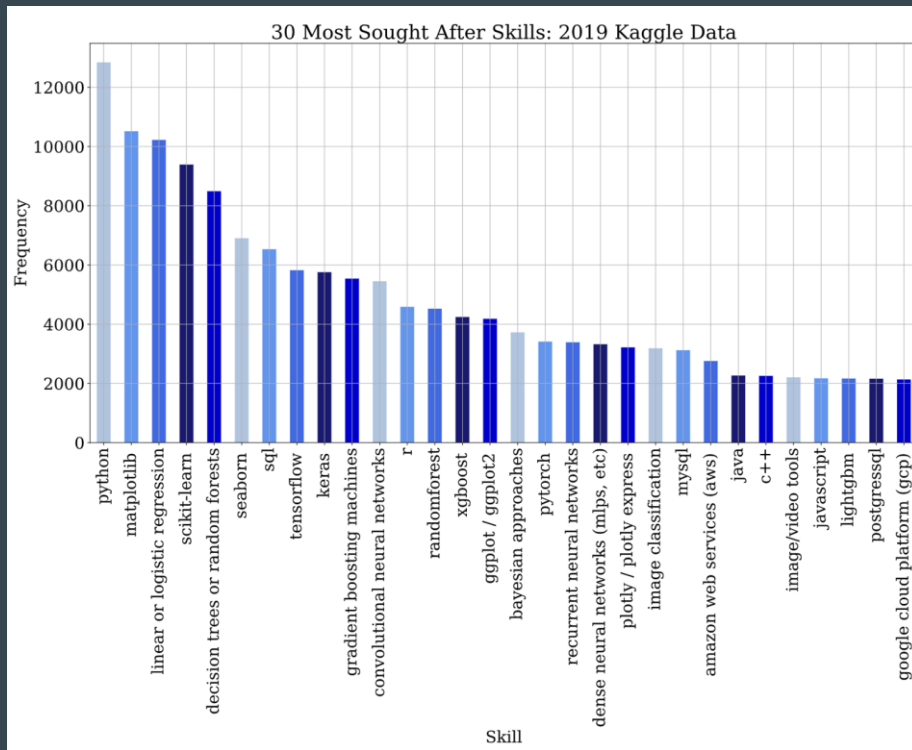
Most Used/Desired Skills

Data Sources:

- 2018/2019 Kaggle data
- Indeed.ca job description data

- Visualized the most commonly used/desired skills from the Kaggle/Indeed data
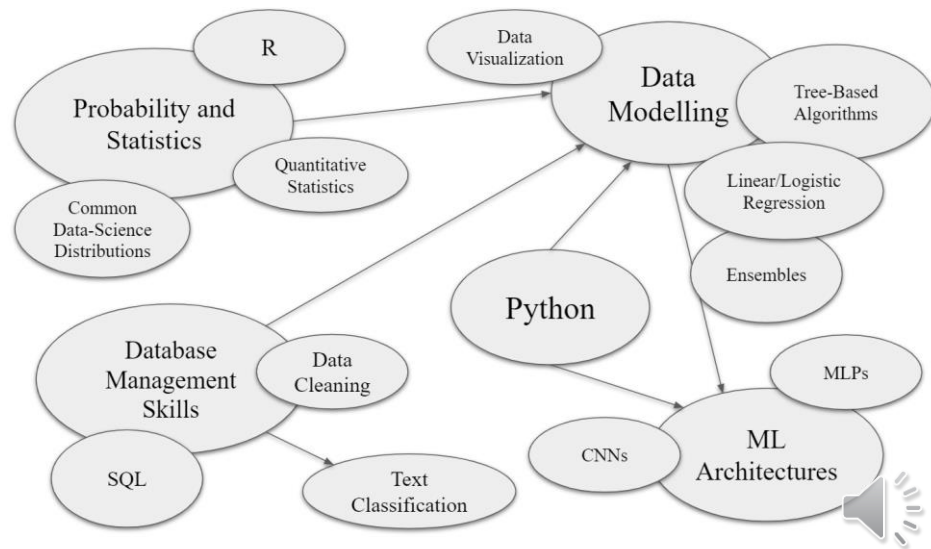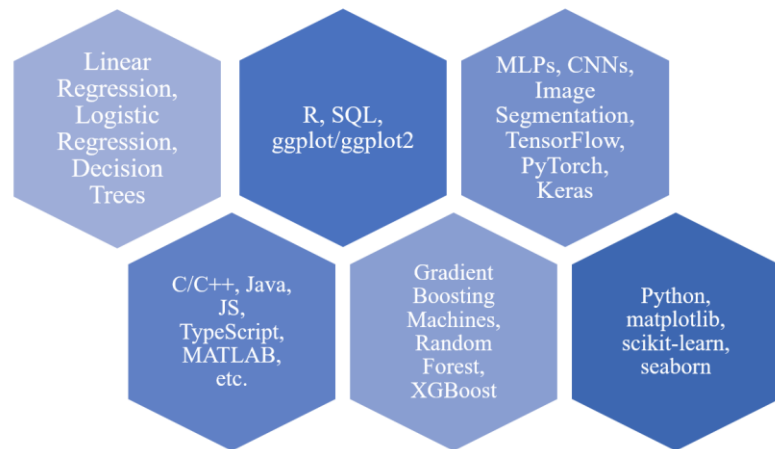
Problems:

- Skills are unorganized
- Skill dependencies exist

30 Most Sought After Skills: 2019 Kaggle Data

Skill Clustering and Dependency Graph

- Skill clustered using agglomerative clustering
- Dependencies modelled using directed graph

Curriculum Design

Lectures 1-3
- ➔ DB management skills, data cleaning, SQL
- ➔ Common distributions, sampling techniques, R, ggplot
- ➔ Data visualization and exploratory analysis using Python

Lectures 4-5
- ➤ Loss functions, performance metrics, ethical modelling
- ➤ Fundamental modelling architectures
  - ○ Linear/logistic regression using scikit-learn

Lectures 6-7
- ➤ Decision trees and ensembles
- ➤ Advanced tree-based algorithms
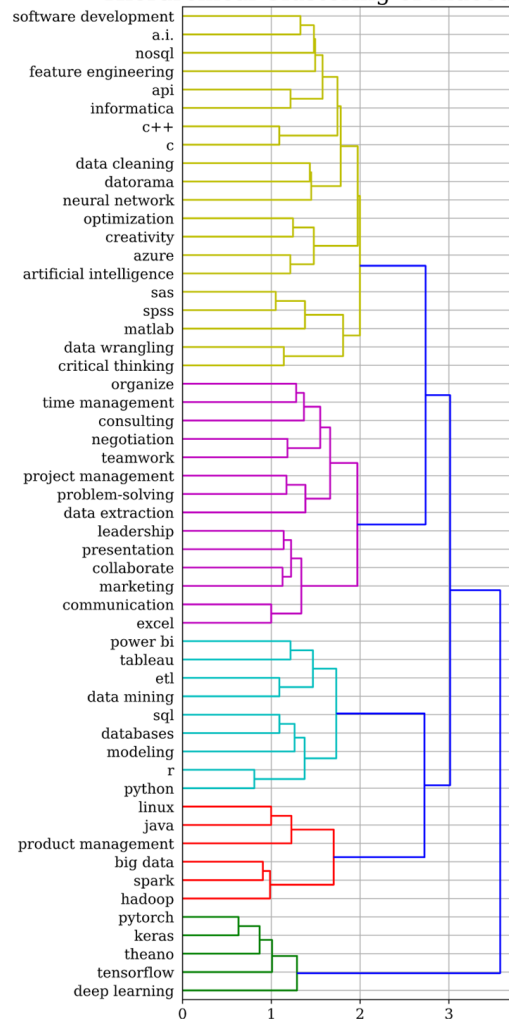- ➤ Introduction to machine learning

Lectures 8-10
- ➤ Dense neural network architectures
- ➤ Image classification using CNNs
- ➤ NLP: BoW, TF-IDF, introduction to learned representations

UofT Master of Data Science and Artificial Intelligence

- Both soft and technical skills are important to employers
    - Communication, teamwork, excel, time management, etc.
- Cluster soft and technical skill groups present in job description data

Hierarchical Clustering of Indeed Skills

Data Science and AI Master's at UofT

Technical Skills

Providing students with the technical skills that they need to succeed in the field of data science

Technical Program Elements
- Lectures
- Evaluations
- Course-work

Course Structure:
➔ *1 foundational course, 6 core courses, 1 elective, and a mandatory internship semester*

Practical Skills

Fostering a professional academic environment that motivates students to step outside of their comfort zone and push their limits

Practical Program Elements
- Mandatory 600 hour internship
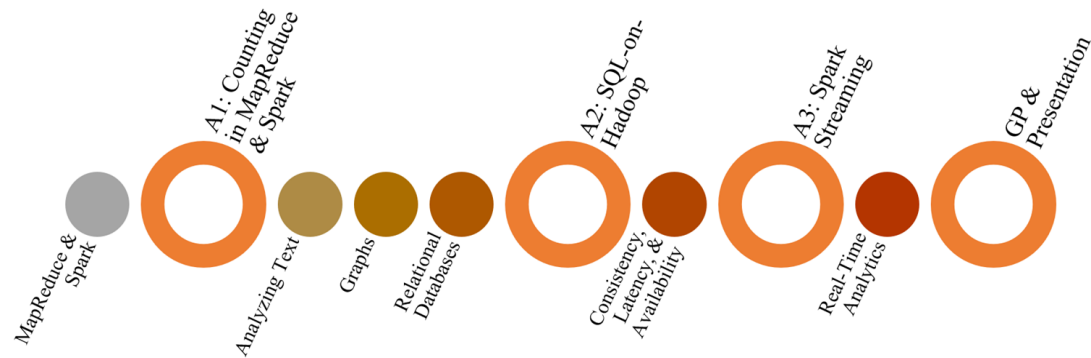- Course assignments, projects, and presentations

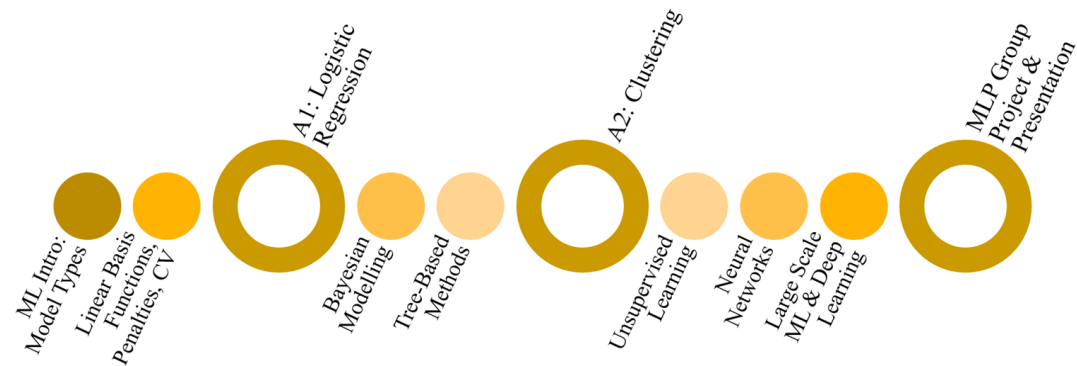# "Master of Data Science and Artificial Intelligence" Program Structure

S1	Internship Semester	S2
- Foundations - Data Mining & Exploratory Analysis - Statistical Learning - Data Visualization		- Introduction to Data Management - Distributed Computing & Big Data - Introduction to AI - Elective

Distributed Computing & Big Data:

- MapReduce & Spark
- A1: Counting in MapReduce & Spark
- Analyzing Text
- Graphs
- Relational Databases
- A2: SQL-on-Hadoop
- Consistency, Latency, & Availability
- A3: Spark Streaming
- Real-Time Analytics
- GP & Presentation

Statistical Learning:

- ML Intro: Model Types
- Linear Basis Functions, Penalties, CV
- A1: Logistic Regression
- Bayesian Modelling
- Tree-Based Methods
- A2: Clustering
- Unsupervised Learning
- Neural Networks
- Large Scale ML & Deep Learning
- MLP Group Project & Presentation

Core Course Visualization

Identifying Domestic Data Science Opportunities

Company Evaluation Metrics:

- Work environment
    - Work/life balance
    - Culture
    - Career opportunities
- Compensation

Work Environment Data:

Rating Data on Glassdoor
- Culture & Values, Work/Life Balance, Overall Rating
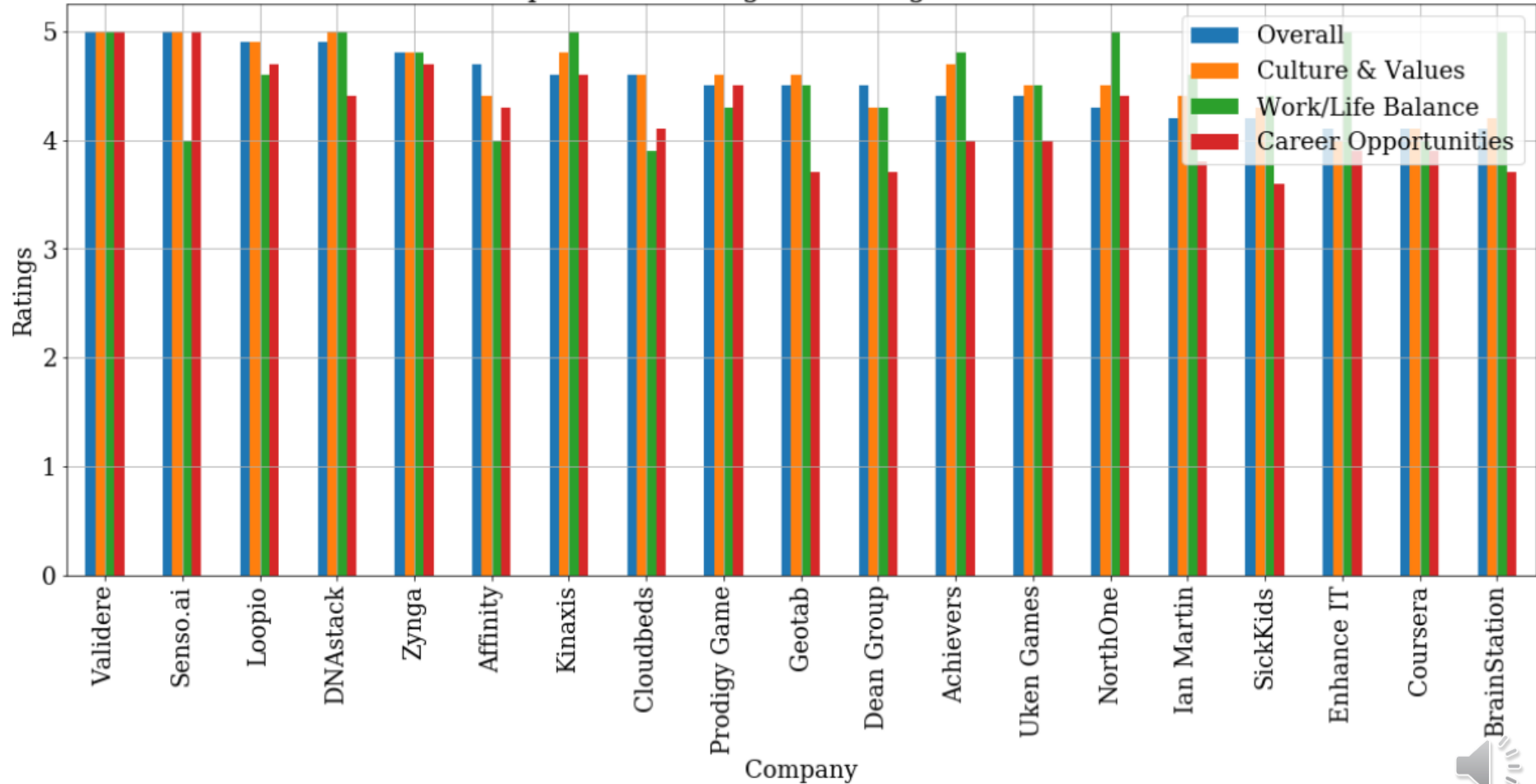
Compensation Data:

Salary Data on Workopolis

Company Evaluation: Work Environment

- Scraped job review data from Glassdoor
  - Culture/Value Rating
  - Work/Life Balance Rating
  - Career Opportunities Rating
  - Overall Rating
- Top companies determined by thresholding on Overall Rating and Career/Opportunities Rating

Companies With Highest Ratings on Glassdoor
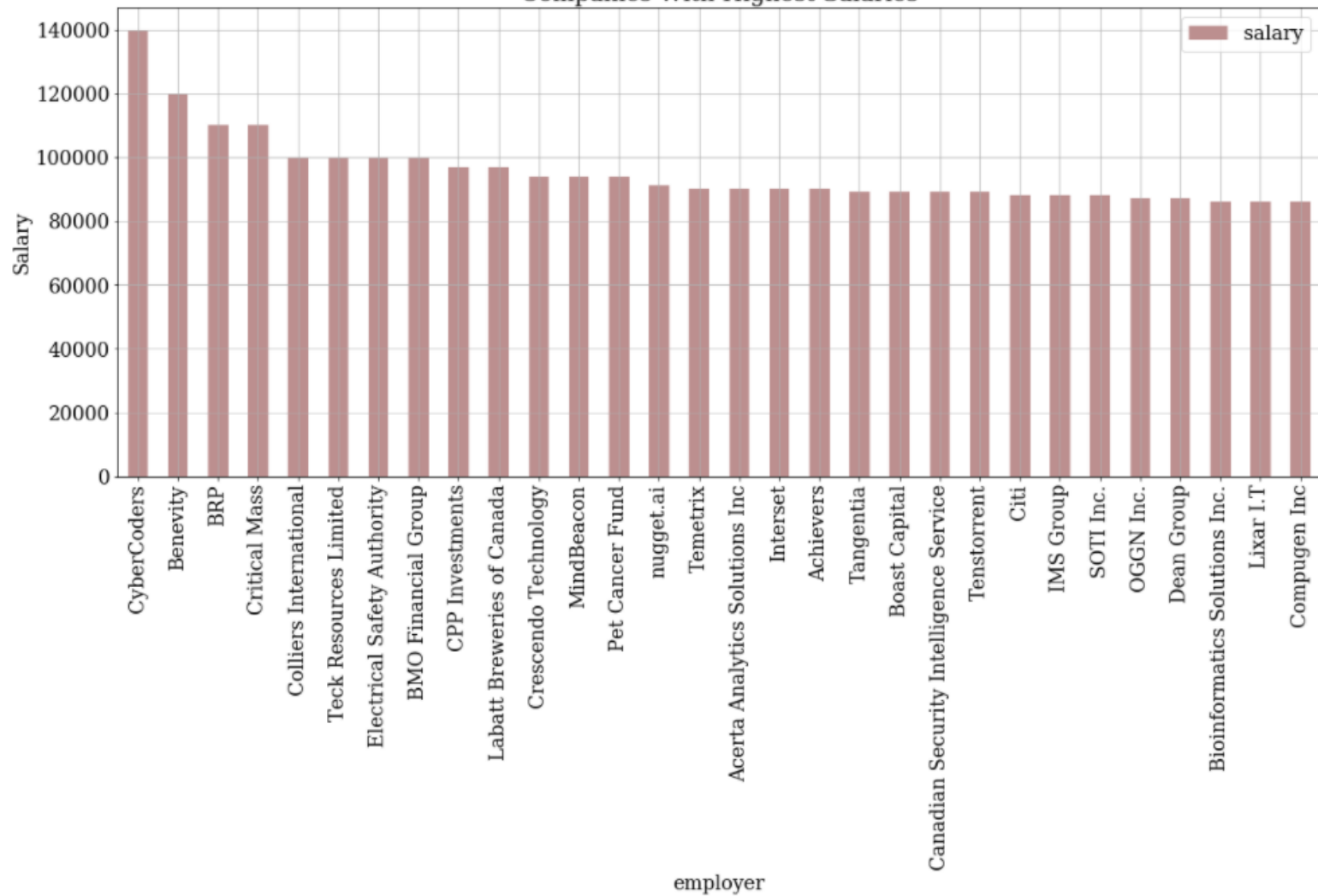
Company Evaluation: Compensation

Scraped salary data from Workopolis

Issues:

- Average employee may not be represented by people leaving reviews
- We are implicitly assuming that the data science opportunities at different companies deserve similar pay
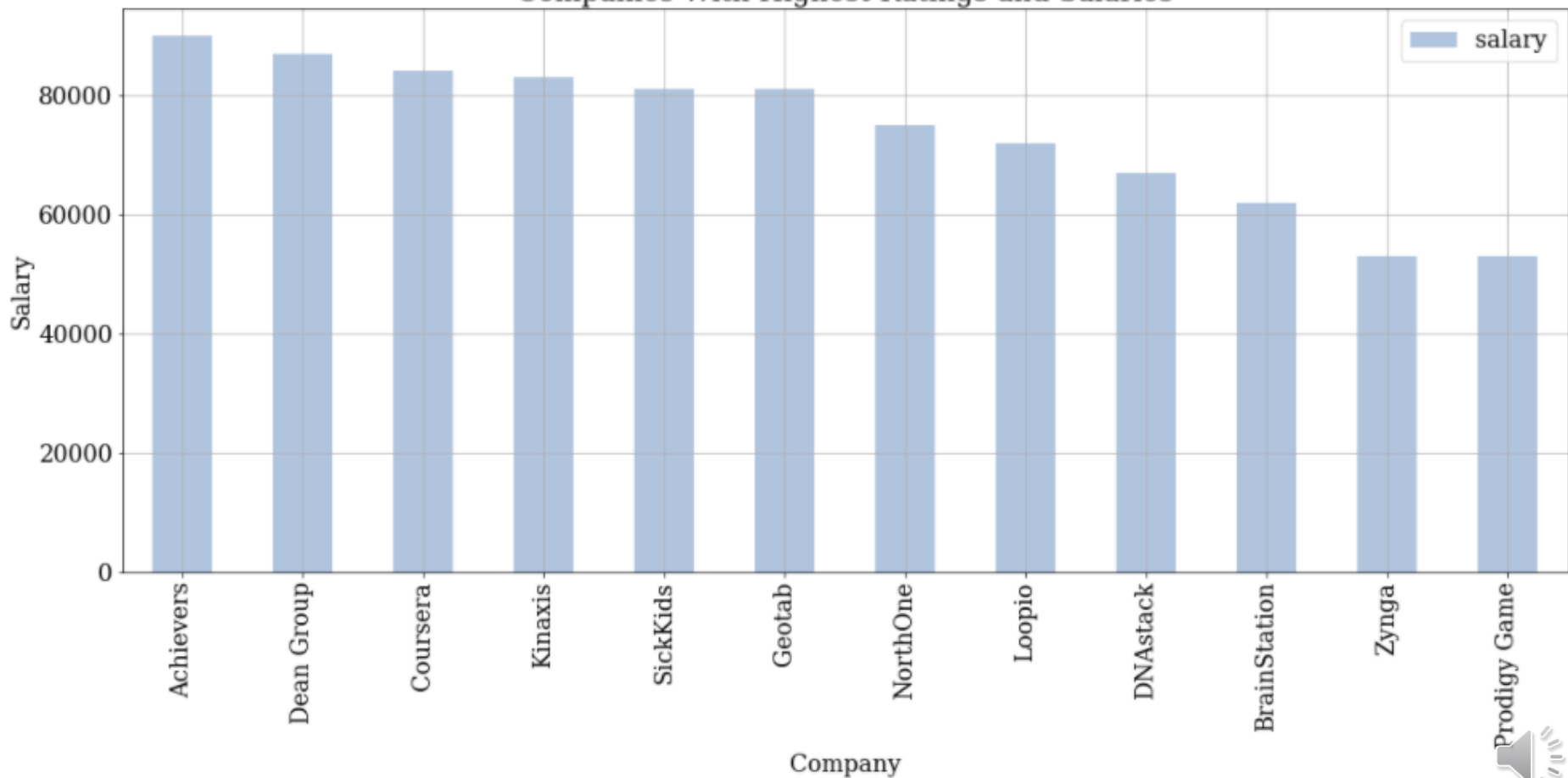
Companies With Highest Salaries

Company Recommendations

- Combining review data from Glassdoor and salary data from Workopolis
    - Highest-rated companies thresholded on our salary metrics

Companies With Highest Ratings and Salaries

Recommendations

Achievers

Employee voice and recognition solution. Industry leader in engagement.

Dean Group

One of the top Canadian-based employment agencies.

Coursera

Offers personalized course recommendations and top-tier online education.

Kinaxis

Revolutionary supply chain agility for supply chains around the world.