# Objective

As the company is rapidly expanding, increasing contract win rate and acquiring new clients have been the key focus.

My goal as the data scientist is to ensure company's competitive advantage of our technology and scientific solutions in the hiring market.

In order to determine the a compensation package that meets the market need, we have to first identify the industry factors that are most important in predicting the salary amount.

# Data collection

To understand more about the current market trend on data related jobs, I have scraped data from "My Career Future" ([https://www.mycareersfuture.sg](https://www.mycareersfuture.sg/)). Keywords that are used for search query are 'analyst','data engineer','data scientist','data analyst','research scientist' and 'business intelligence'. The job postings include information such as job title, industry, company info, job requirement, seniority, range of expected salary, industry and employment type.

A total of 1,593 job postings have been collected.

Of all the job postings scraped:

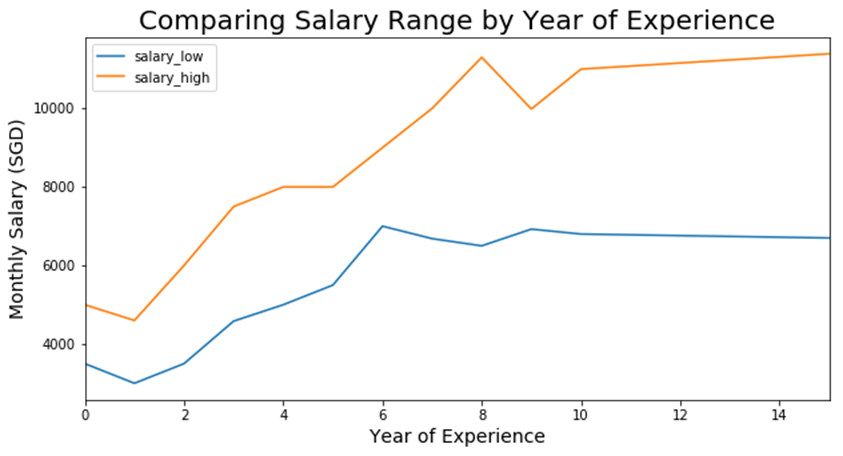
* 1,218 jobs with analyst title
* 122 jobs with data engineer title
* 106 jobs with data engineer title
* 316 jobs with other title but data related job scope

# Key findings

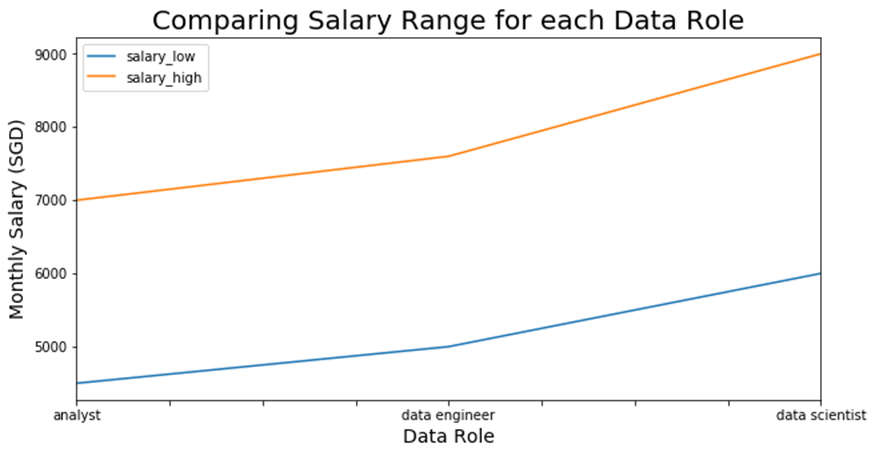
## High level summary

On average, the salary offered in the market is SGD 54K.

50% of the jobs are looking for experienced candidates with a minimum of 2-3 years relevant experience.

Average monthly salary to offer for candidates with 2 years experience are ranging from SGD 4.5k to SGD 7k, whereas for candidates with 3 years experience are ranging from SGD 7k to SGD 11k.

## Title categories

Further drill down to job category by job title, we also observe that data scientists are expecting a higher package than data engineers followed by data analyst.

## Companies

Also, a few companies are having a significant number of job openings for data related roles in the past month, which could be a good indicator of data driven culture in a company:

* DBS Bank: 62 job openings
* Machspeed Human Resources: 32 job openings
* Kerry Consulting: 27
* GovTech: 27
* Bluechip Platforms: 26
* OCBC Bank: 25
* A-star Research Entities: 25

Etc

## Industries

The top 5 industries which offer the highest package in data related roles are:

* Consulting, Marketing/Public Relations, Risk Management, Insurance and Advertising/Media.

However, the top 5 industries which have the highest demand for data-skilled individuals are:

* Information Technology, Accounting/Auditing/Taxation, Engineering, Consulting and Sciences/Lab/R&D.

## Job descriptions

I've also performed Natural Language Processing (NLP) analysis to investigate if the job requirements impact salaries. A quick glance on the most common used keywords in the job requirements based on different roles are:

* Data Scientists:
  + Machine Learning
  + Data Science
  + Big Data
* Data Engineers:
  + Big Data
  + Hands-on Experience
  + Machine Learning
* Analysts:
  + Communication Skills
  + Business Requirements
  + Problem Solving

It's obvious that organizations are focus more on interpersonal skills (soft skills) when looking for analysts as opposed to data scientists or data engineers which require a higher level of technical skillsets (hard skills).

# Predicting Minimum Expected Salary

## Factors used for prediction

By using information mentioned above, I've run a few predictive models that will estimate minimum annual salary to offer based on:

* Basic job postings information:
  + Minimum experience required
  + If the job requires PHD
  + Title category ie data scientists vs data engineers vs analysts
  + Industry
  + Seniority
  + Employment Type ie Full time vs Contract
  + Company name
* Technical skill sets required for the job ie Python, Hadoop, Spark, Tableau, AWS etc.
* Keywords extracted from job requirements and company description

## Identifying key factors impacting salary

A few models are evaluated by looking at the median absolute error. The best model is suggesting that the error in estimation is about SGD 9.5k in annual salary, which is about SGD 790/ month.

For interpretation purpose, I'll pick the 2nd best model which allows me to not just explain the magnitude but also the direction on how the factor is impacting the salary.

The 2nd best model is suggesting a median error of SGD 10k, which is about SGD 833 / month.

The most important factors identified from the model in impacting the salary are:

1. Minimum year of experience required
   * An estimated amount of SGD 3.7k increase to your annual salary with an increase in one year of relevant experience, bench-marking against average pay of SGD 58k in the market.
   * That suggests 6% increase yoy with an increasing experience.
2. Seniority
   * For job postings with a seniority level of Junior Executive, Fresh/Entry Level or Executive, the model is suggesting a decrease of SGD 2k to SGD 3k on annual salary, bench-marking against average pay of SGD 58k in the market.
3. Companies
   * A few companies are found to be able to offer a more attractive package as compared to roles with the same level of experience and job scope for example:
     + Tangspac Consulting
     + Lazada South East Asia
     + Barclays
     + Twitter
     + Standard Chartered Bank
     + Google

## Top over/under estimated jobs

The model tends to over-estimate for jobs that offer below-mean salary and under-estimate for jobs that offer above-mean salary.

Take a closer look to:

1. Top 3 over-estimated jobs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Starhub | Technical Support Analyst (contract) | Contract | 20400 | 46690 |
| Dathena Science Pte Ltd | Field Data Scientist | FullTime, Internship | 24000 | 57882 |
| Sun Electric (Singapore) Pte Ltd | Pricing Manager | FullTime | 19200 | 61449 |

1. Top 3 under-estimated jobs:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DRW Singapore Pte Ltd | Trading Analyst | FullTime | 84000 | 41097 |
| Hartree Partners Singapore Pte Ltd | Analyst | FullTime | 84000 | 42349 |
| OCBC | Business Analyst Murex, Global Treasury (Business Process Transformation, 2 year contract) | Contract, FullTime | 84000 | 42526 |

## Limitations and future development

* Minimum education level required needs to be included as one of the factors however it does not come as one of the individual data points in each job posting. It can potentially extracted from the job description but it requires intense effort of data cleansing.
* More than 30% of the jobs have no minimum experience requirement. The missing values are now filled with the median however, a more accurate estimation of the year of experience is required to better predict the salary.