# DATA SCIENCE Jobs Analysis Project.

A Moringa Group 4  Project

# TABLE OF CONTENTS

# 01

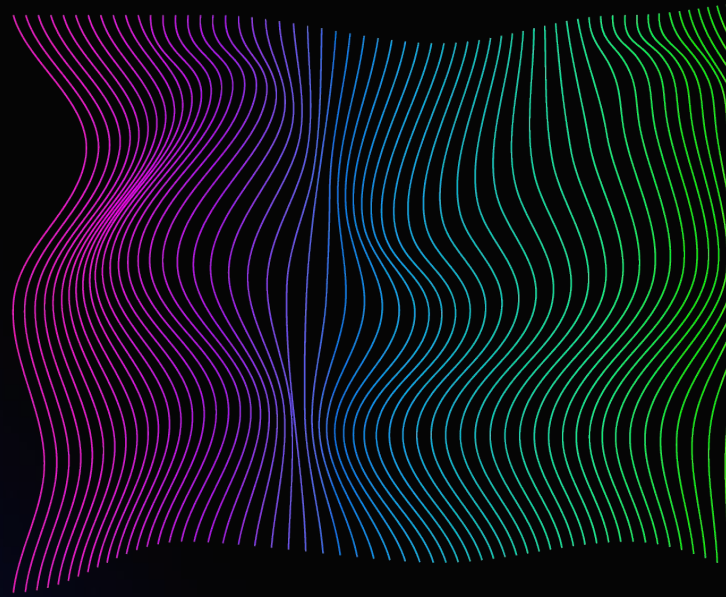
# Introduction.

# KenyaData Insights

KenyaData Insights is a forward-thinking data company in Kenya that specializes in comprehending data science salary expectations. The business is dedicated to examining how these expectations affect worker engagement and happiness at every level of the organizational hierarchy.

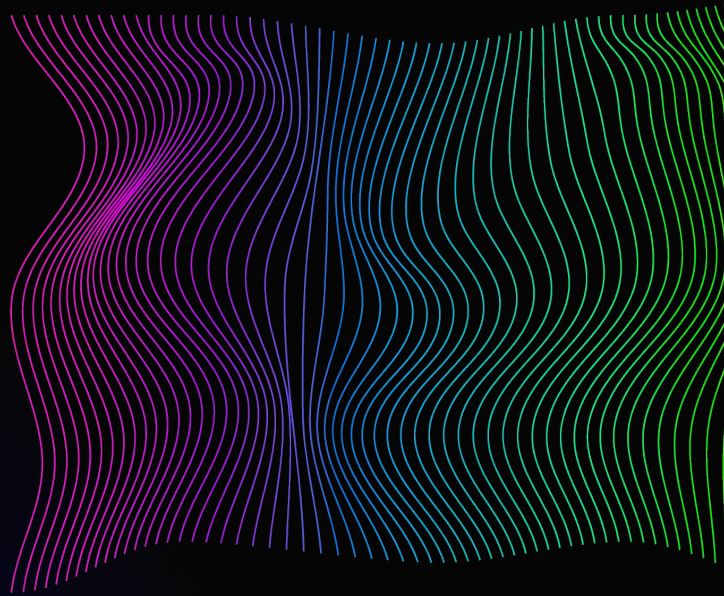
# Business Problem

KenyaData Insights makes it a priority to develop a predictive model to analyse and forecast salaries for data science job positions based on factors such as job titles, experience level, company size, and other relevant variables.

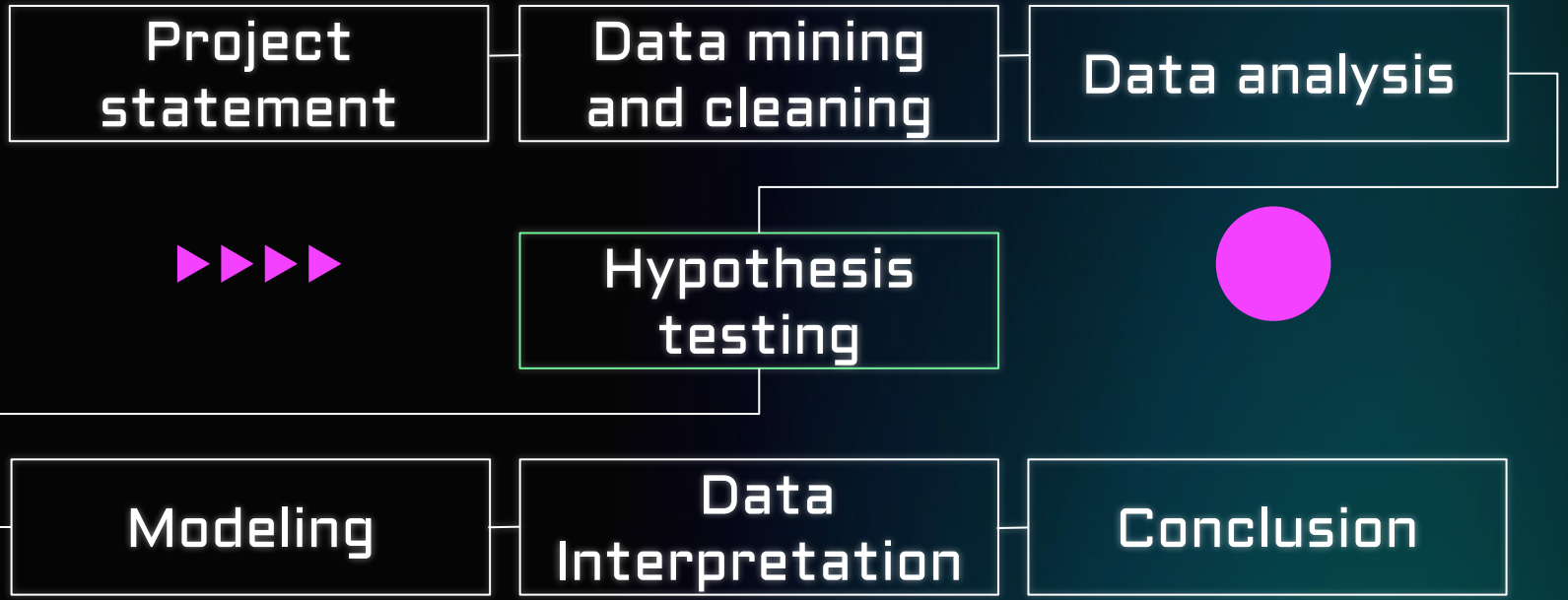


# Overview

The goal of the study is to better understand the complex relationship that exist within the data and machine learning business. Specifically, the study will look at how job features, compensation structures, and the dynamics of remote work interact with one another.



# DATA PROJECT ARCHITECTURE

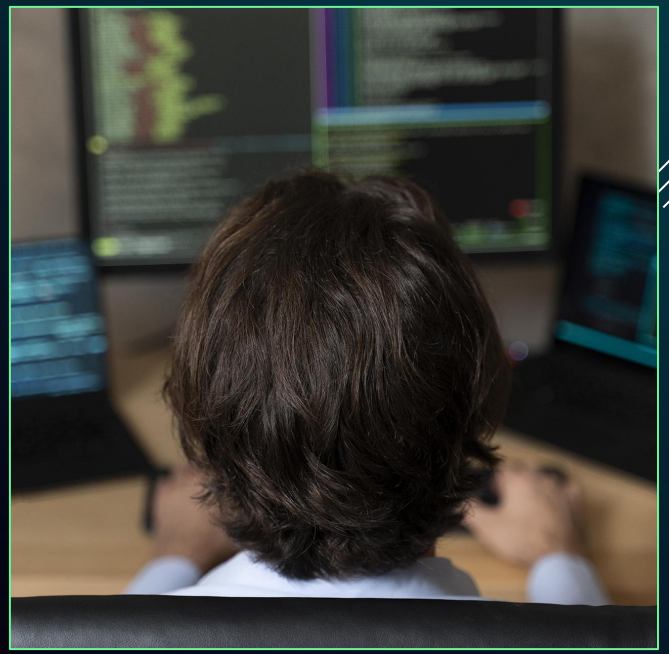


# Objectives

02

# BUSINESS OBJECTIVES

- Develop a strong salary prediction model
- Optimize placement strategies for future data scientists
- Improve recruitment processes with data-driven insights
- Build ethical data governance standards
- Cultivate engagement with educational institutions and business partners
- Ensure responsible data usage
- Improve alignment between educational and industrial demands



# STAKEHOLDERS

- Investors
- CEOs
- Board of Directors
- Competitors
- Advisors
- Consultants
- Job seekers

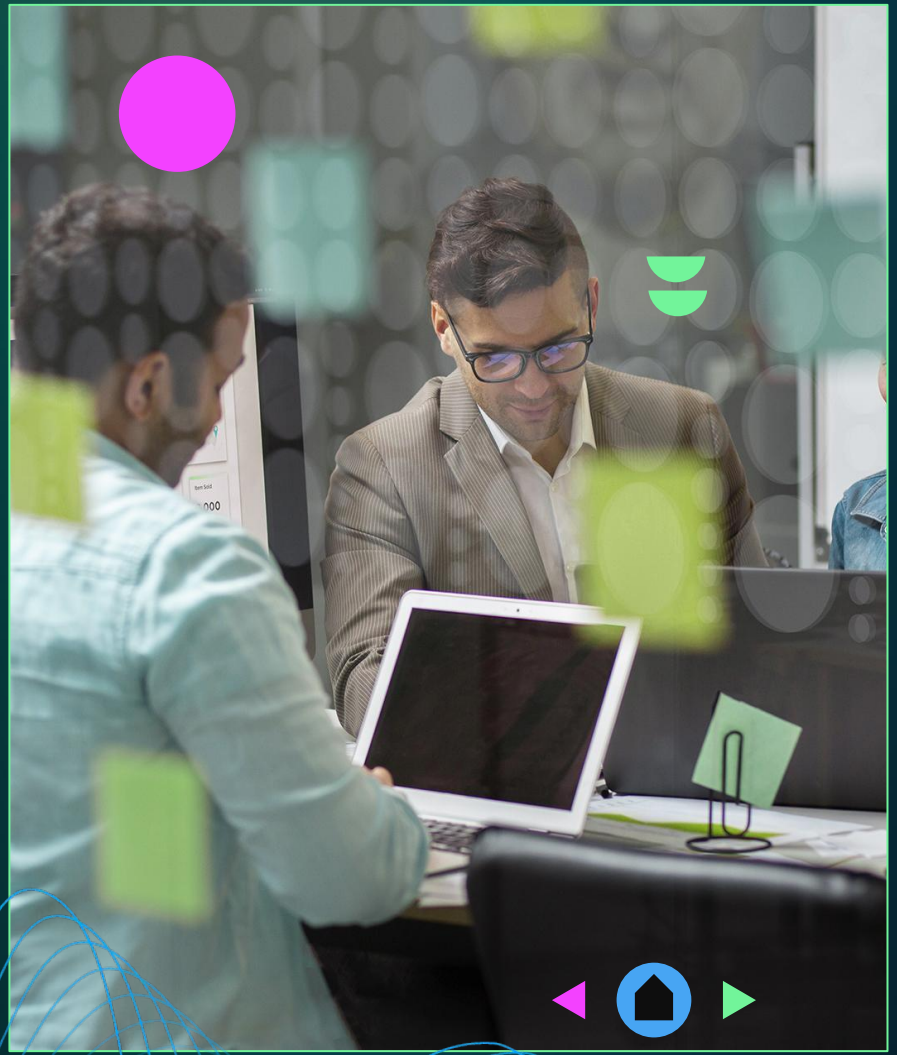
# 03

# Data Understanding.

# DATA SOURCE

https://www.kaggle.com/code/hasibalmuzdadid/data-science-jobs-salary-analysis-retro-vibe/input

Our dataset, Data science jobs salary analysis, was collected from Kaggle. It contained 607 rows of different job listings with corresponding information such as experience level, employment type, job title, employee residence, company location, company size, salary etc within the work years 2020 -2022. After performing hypothesis testing on a number of variables, we settled on the three that seemed to have a significant impact on salary earned.



# HYPOTHESIS

i.) Job Titles

- Null: There is no difference in compensation across various job titles
- Alternative: Compensation levels vary significantly based on different job titles.

ii.) Work Experience

- Null: Work experience does not have a significant impact on salary levels.
- Alternative: Employees with longer work experience receive higher salaries.

iii.) Company Size  Salary Disparities:

- Null: There are no significant differences in salary distributions across various company sizes
- Alternative:  Salary distributions vary significantly across different company sizes

# FACTORS AFFECTING COMPENSATION

Regions

Job Title

Work Experience

# 04

# Data Analysis

# DATA PREPARATION

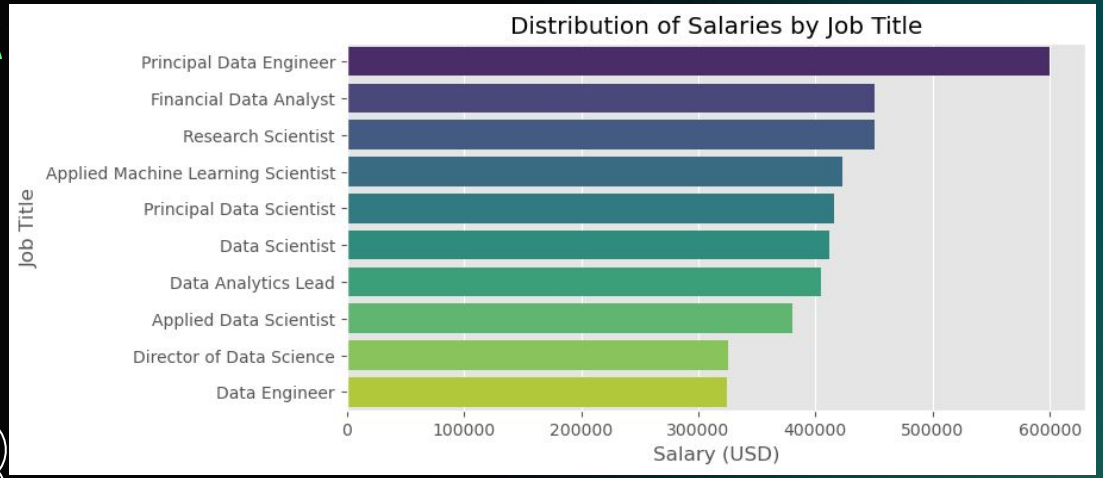
We took some preprocessing steps to ensure we had data that was easy to work with:

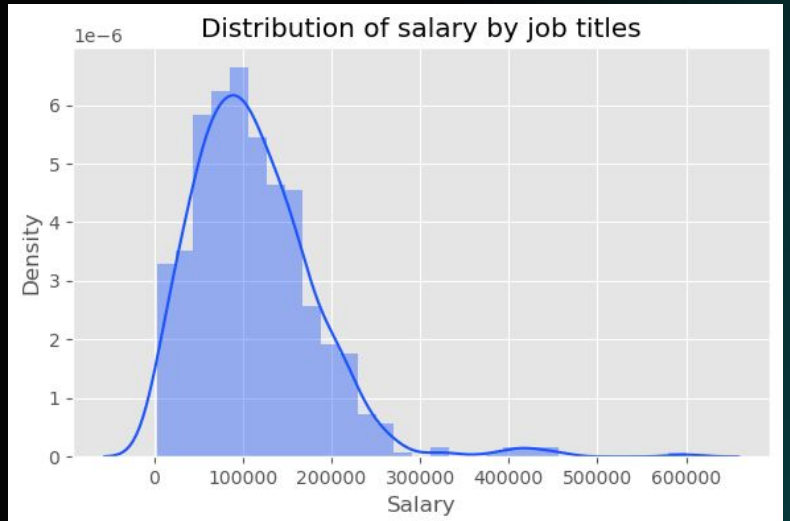
- We chose to work with a standardised salary currency i.e USD
- We also peeked into the top 10 job titles and salary ranges to get an overview of the data.
- In doing so, we realised that there was an outlier that greatly impacted our data. We chose to retain it in the data set but use the mean when doing the regression model.
- We then went ahead and dropped columns that would not be viable for our regression model.

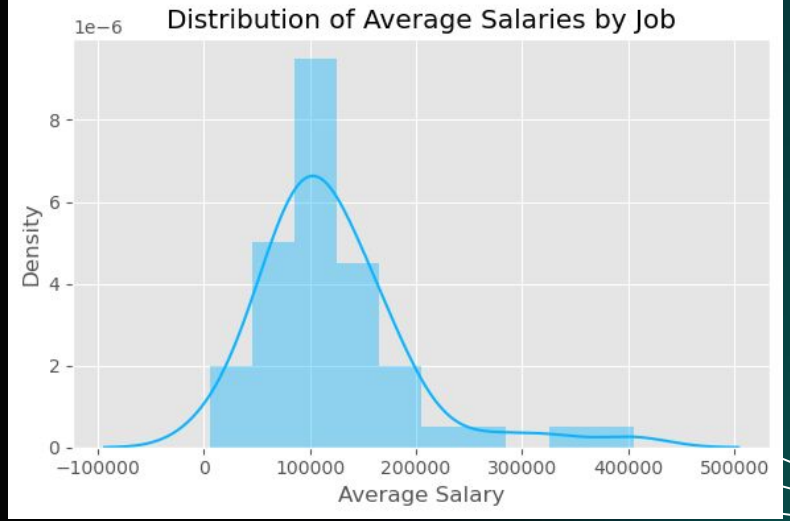
# 05

# Data Visualisation.

# Top job titles and salaries



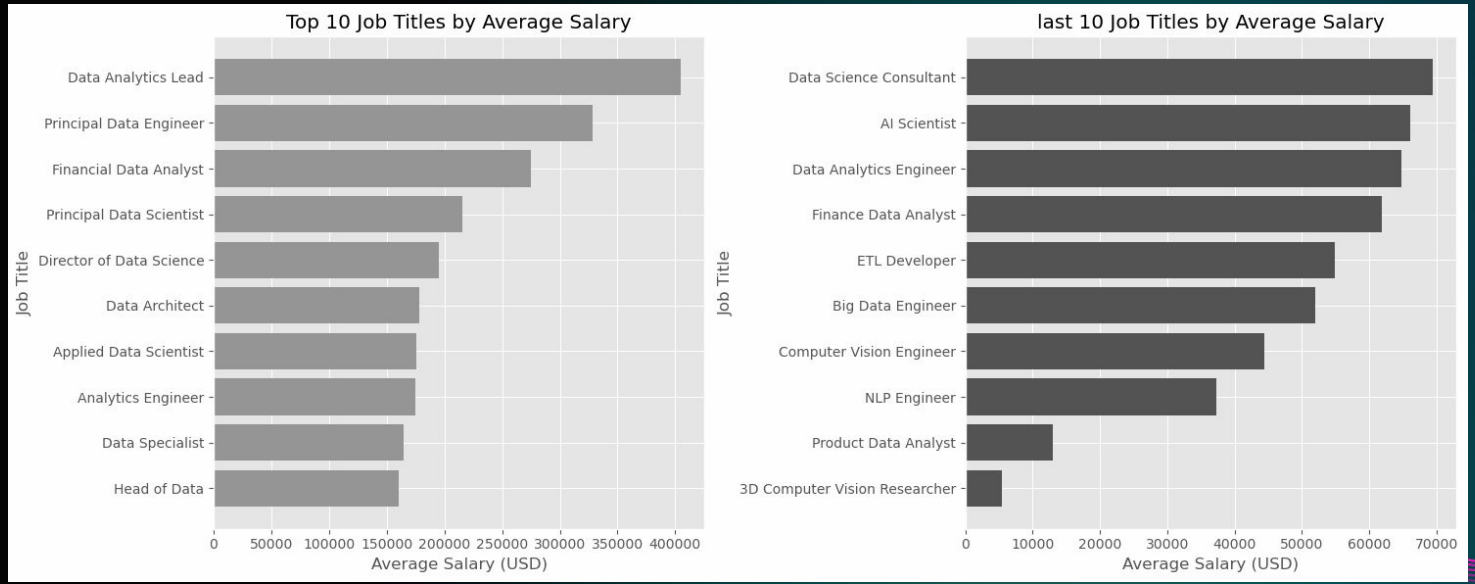



- A look at the top 10 job titles against salaries reveals that the average salary range per year in USD is 200,000.
- We however notice that the highest paid is way out of the average range skewing our density plot to the right

- Below it we have a distribution plot of average salaries that is more representative of the larger population

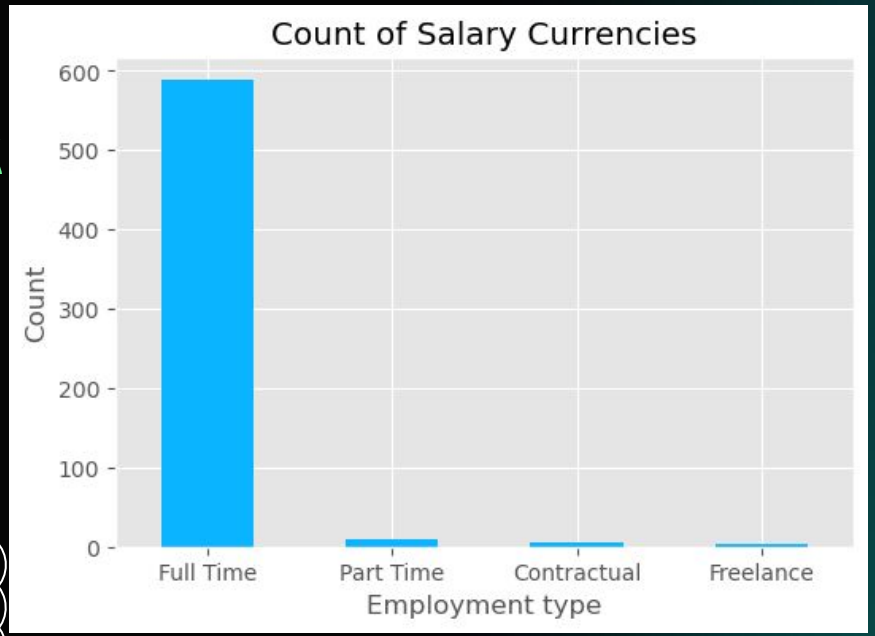
# Comparison of top and bottom ten job titles

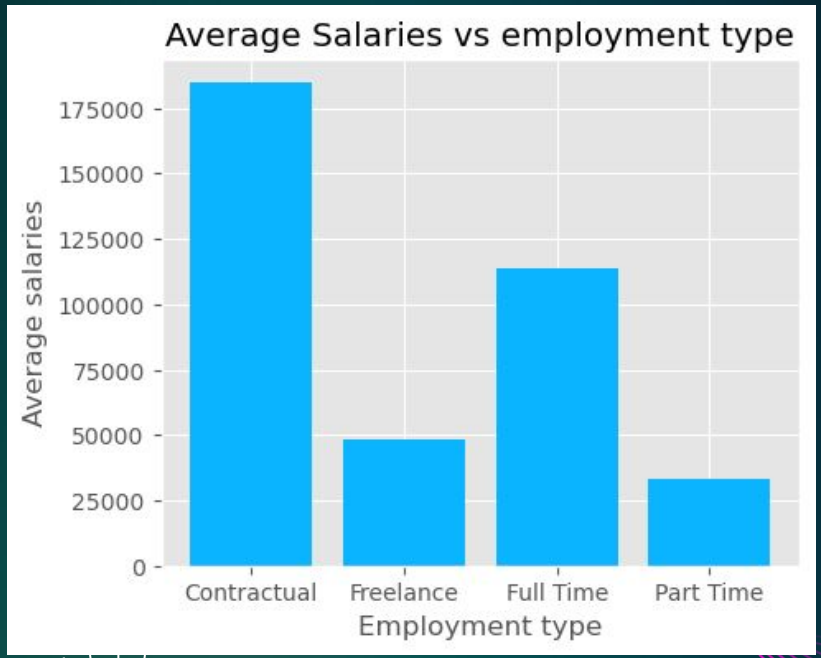


# $412,000

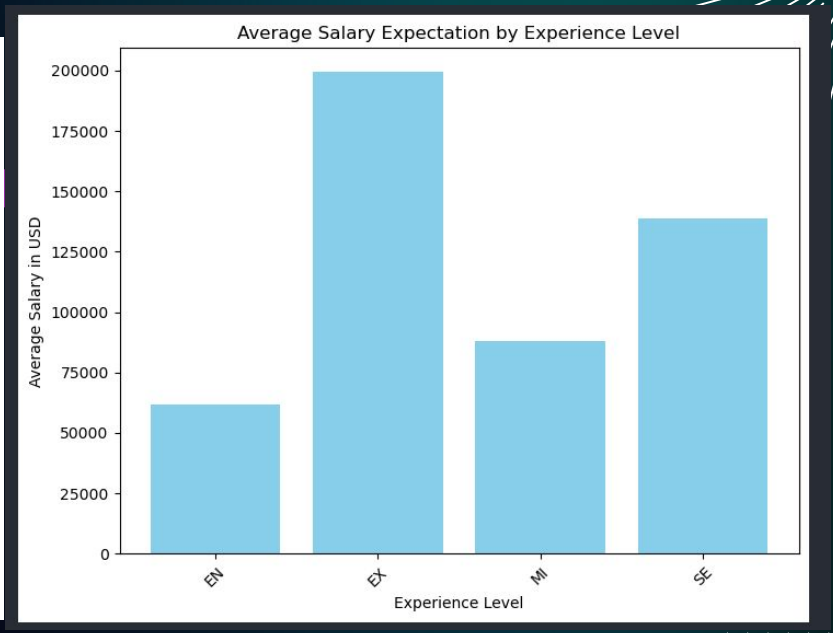
It is interesting to note that, data scientist is not among the top jobs as we had expected. However, here is the average annual salary

# Employment type against salaries



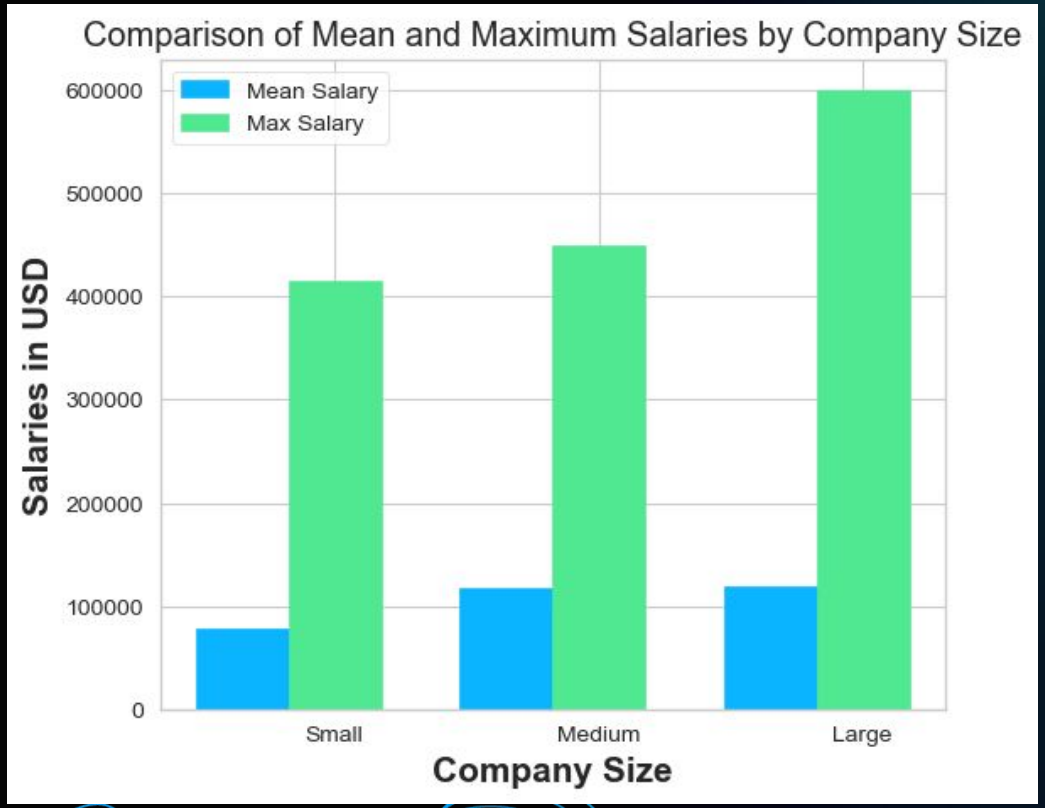




While full time employment is the most prevalent, contractual employment is the highest earner


The most common level of experience is the Senior Level while the least being the experienced

As is seen here, The highest paid is the experienced level followed by Senior level as expected




From this bar graph, we can observe that smaller companies pay less in general and larger companies pay more.

# REGRESSION ANALYSIS

We dropped columns that are not metric

We made a simple linear regression using salary in US Dollars against salary expectations that showed that the model is statistically significant

Hot encoded experience level, employment type and company size data in order to do a multiple linear regression

# REGRESSION INTERPRETATION

Using large company size as our variable, there is an associated decrease in salary scaled at 0.0193 when the company size is medium and a further 0.0475 when the company is small

Using entry level experience as our reference variable, there is a corresponding increase in salary scaled at 0.2 for experts, 0.03 for mid-level and 0.12 for senior level employees in relation to entry level

In this case, our reference variable is contractual employment. Associated with it is a decrease in salary level of 0.21 for freelancer, 0.12 for full-time employees and 0.18 for part-time

# Conclusion

06

# Conclusions & Recommendations:

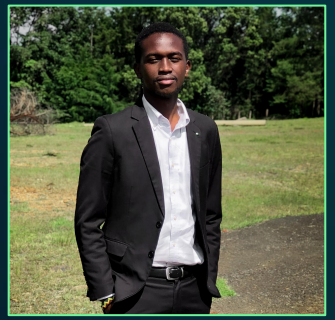
- A full time salary is the most common form employment
- Paying in USD (United States Dollar) is the best form of payment
- Small company CEO's can expect to pay the least in salaries
- Large company CEO's can expect to pay the most in salaries
- Very experienced employees are very expensive <$110,000 USD
- The most common type of Data Science jobs have Senior level expectancy

AVERAGE SALARY EXPECTATIONS

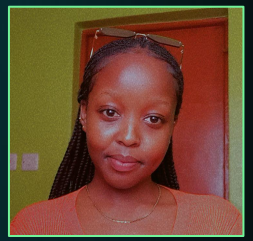
- Entry level experience employees should expect: $USD 61,643
- Mid level experience employees should expect: $USD 87,996
- Senior level experience employees should expect: $USD 138,671

# OUR TEAM

Esther Omulokoli



Edgar Kiprono



Margaret Thiga

# OUR TEAM

Pete Njagi

Jeremy Ngugi

James Koli


THANK
YOU!!