



# Resume Parser Using Nlp & Clustering

BY :

Afnan Alshehri  
Amena Algroon

# TABLE OF CONTENTS

- 01 INTRODUCTION
- 02 DATA ANALYSIS
- 03 TOPIC MODELING
- 04 CLUSTERING
- 05 MODEL



# Backstory :

A resume is a formal document that serves to show a person's career background and skills.

The purpose of this project is to help large companies to scan and filter resumes. So we will use machine learning algorithms (NLP) to help them .

# DATASET



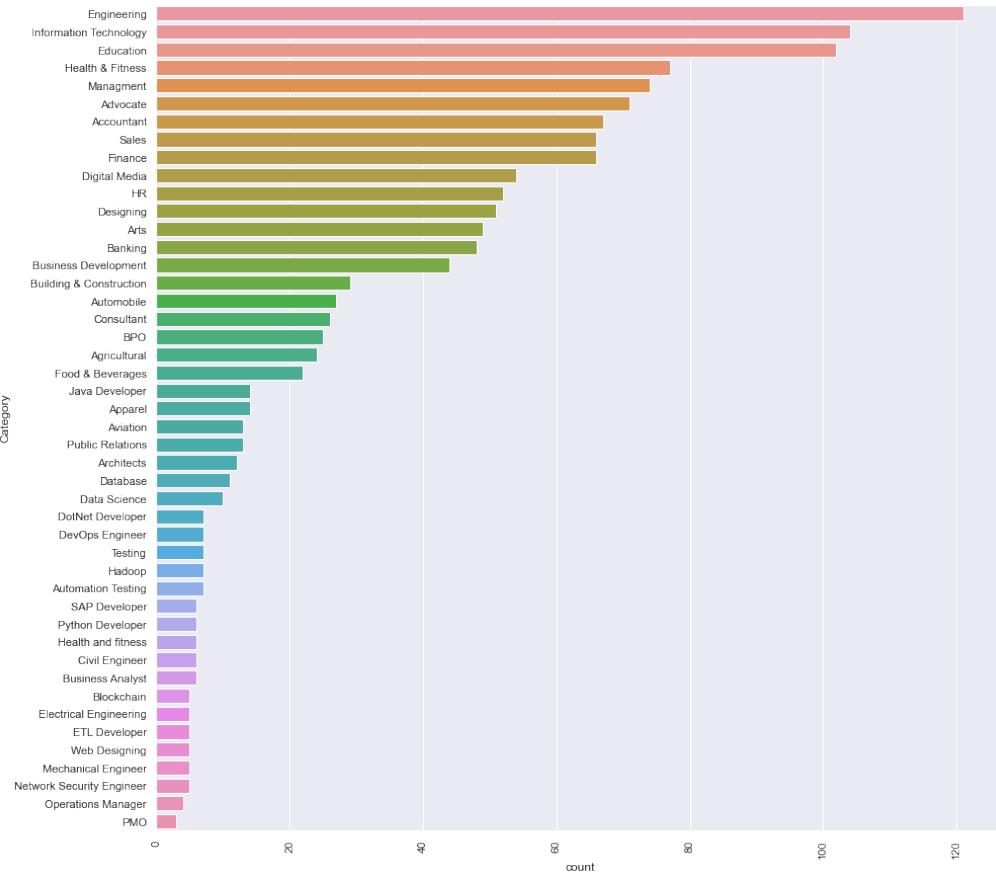
The dataset from Kaggle and github  
Before cleaning 1388 rows and 3 columns  
After cleaning 1297 and 2 columns



# TOOLS



# EDA



# EXPLORATORY ..



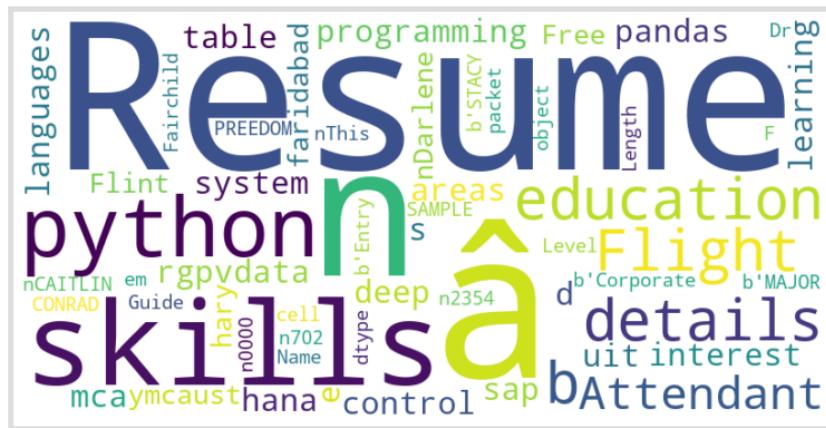
## EXPLORATORY DATA ANALYSIS

- NO Null Values
  - Duplicates (169 )
  - Data Cleaning
  - Vectorization
- Removed Unneeded Column (ID)



# WordCloud Most Common Words

# Before stopword



## After stopword



# TOPIC MODELING

## LSA

### Topic 1

java software developer design application server development

### Topic 2

company customer business sql reports team

## LDA

### Topic 1

web java software developer sql application design

### Topic 2

sales marketing customer product business market service

## NMF

### Topic 1

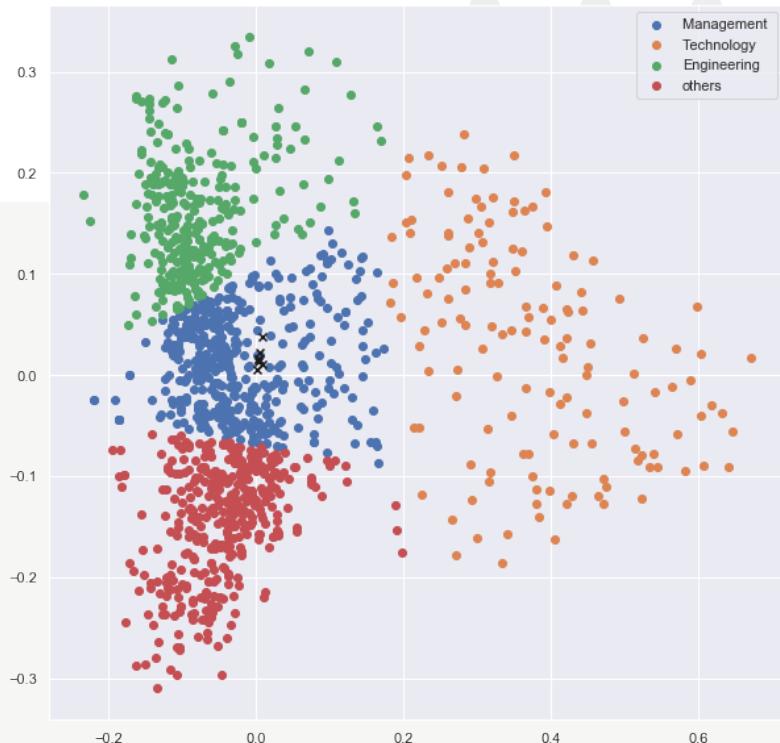
project management business company development system manager

### Topic 2

sales business management market program financial manager



# CLUSTERING:K-mean



## MODEL ( SVM ):

### SVM With Count Vectorizer

Accuracy : 0.5461538461538461

### SVM With TF-IDF Vectorizer

Accuracy : 0.6423076923076924

# Thank You

The End