

Meet our Team



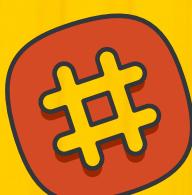
1



2



3



Shriya Arora



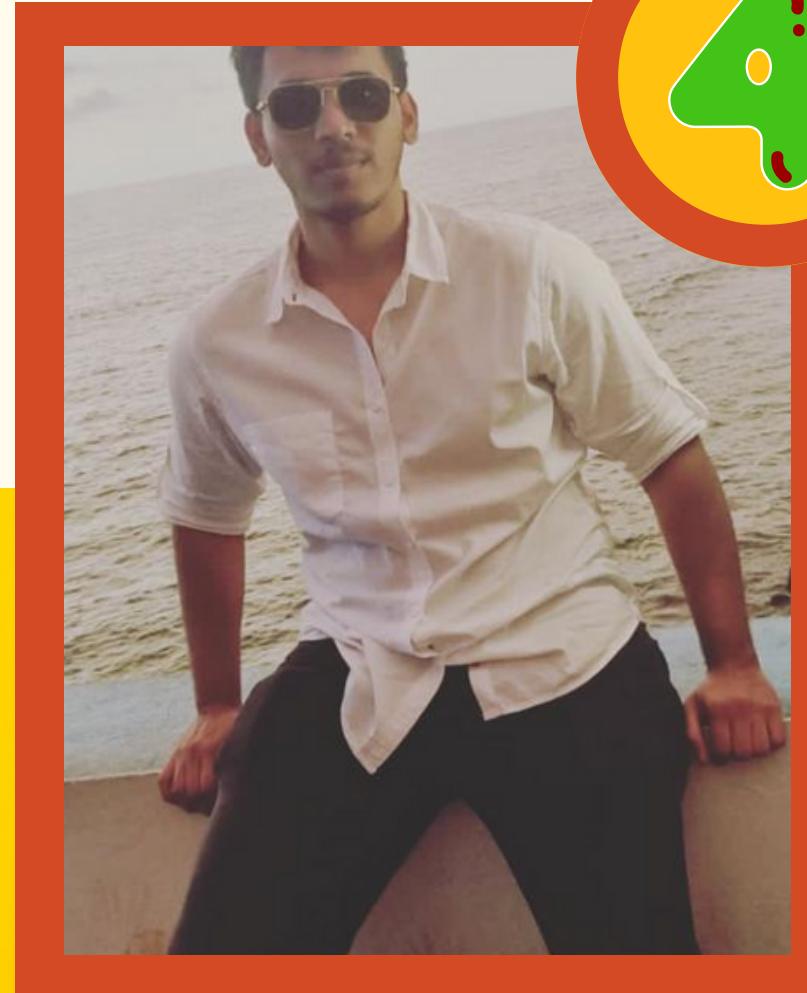
Pallavi Nair



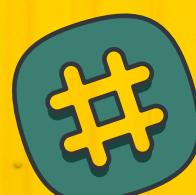
Priyadarshani
Madhusudanan



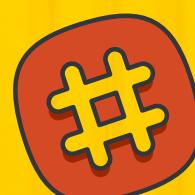
Meet our Team



Aditya Mane



Harshitha Ananthula



Sai Kumar





CSU East Bay

Fake Job Postings

Team 1
BAN 620



Why Fake Jobs?

You & Us



Americans lost \$68 million to job scams this year—here's what to look out for

Published Fri, Jun 10 2022 11:00 AM EDT

April 21, 2021

FBI Warns Cyber Criminals Are Using Fake Job Listings to Target Applicants' Personally Identifiable Information

MONEYWATCH >

Fake job listings are a growing problem in the labor market

MONEY WATCH

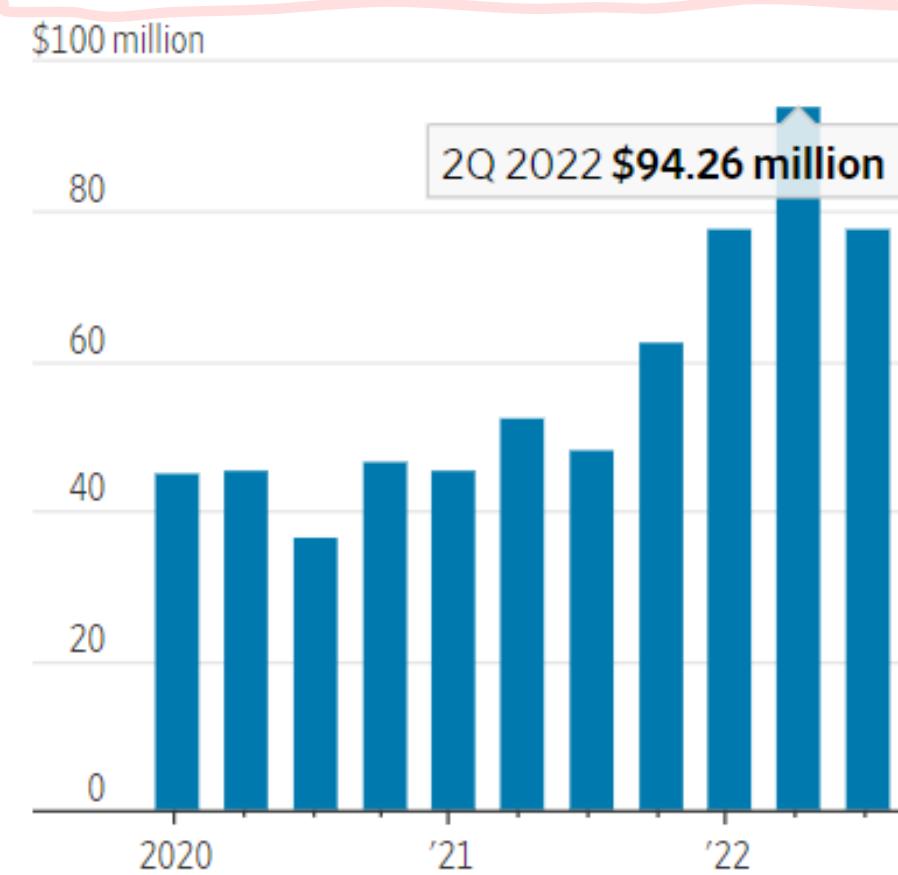
BY IRINA IVANOVA

UPDATED ON: MARCH 30, 2023 / 7:53 PM / MONEYWATCH



Recruiting Scams

How much money Americans lost to fake job proposals, by quarter



Source: Federal Trade Commission

JOB SCAM

Fake Job Scams Are Exploding. Here's How to Catch an HR Impostor

PROJECT

Agenda

Project Overview

What all we did?

Dataset Overview

Explanatory Data Analysis

Model Development & Implementation

Learnings & Moving Forward





Project Overview

We used advanced data analysis techniques to detect and highlight potentially fake job postings. By carefully examining patterns in job descriptions and requirements, our objective was to assist job seekers in avoiding scams and safeguarding their personal and financial information.

PROJECT GOALS

- ◆ Developing a model that can accurately predict fake job postings
- ◆ Using data analysis techniques to identify patterns and trends in fake job postings
- ◆ Raising awareness about the issue of fake job postings



WHAT ALL WE DID

What all We did



Dataset Overview



Source:



Summary: **17,880 observations, 18 features**

Data set consists of job postings from many different countries and has different job positions making it more diverse.

Variable	Description
job id	ID of the job posting
title	Title of the job posting
location	Location of the job
department	Department within the company
salary range	Salary range for the job
company profile	Description of the company
description	Description of the job
requirements	Requirements for the job
benefits	Benefits offered with the job

Variable	Description
telecommuting	Whether telecommuting is allowed
has company logo	Whether the company logo is included
has questions	Whether the job posting includes questions
employment type	Type of employment (full-time, part-time, etc.)
required experience	Required years of experience for the job
required education	Required education for the job
industry	Industry of the company
function	Function of the job within the company
fraudulent (dependent)	Whether the job posting is fraudulent or not





Initial Overview

Figure 1: Missing Values in Features

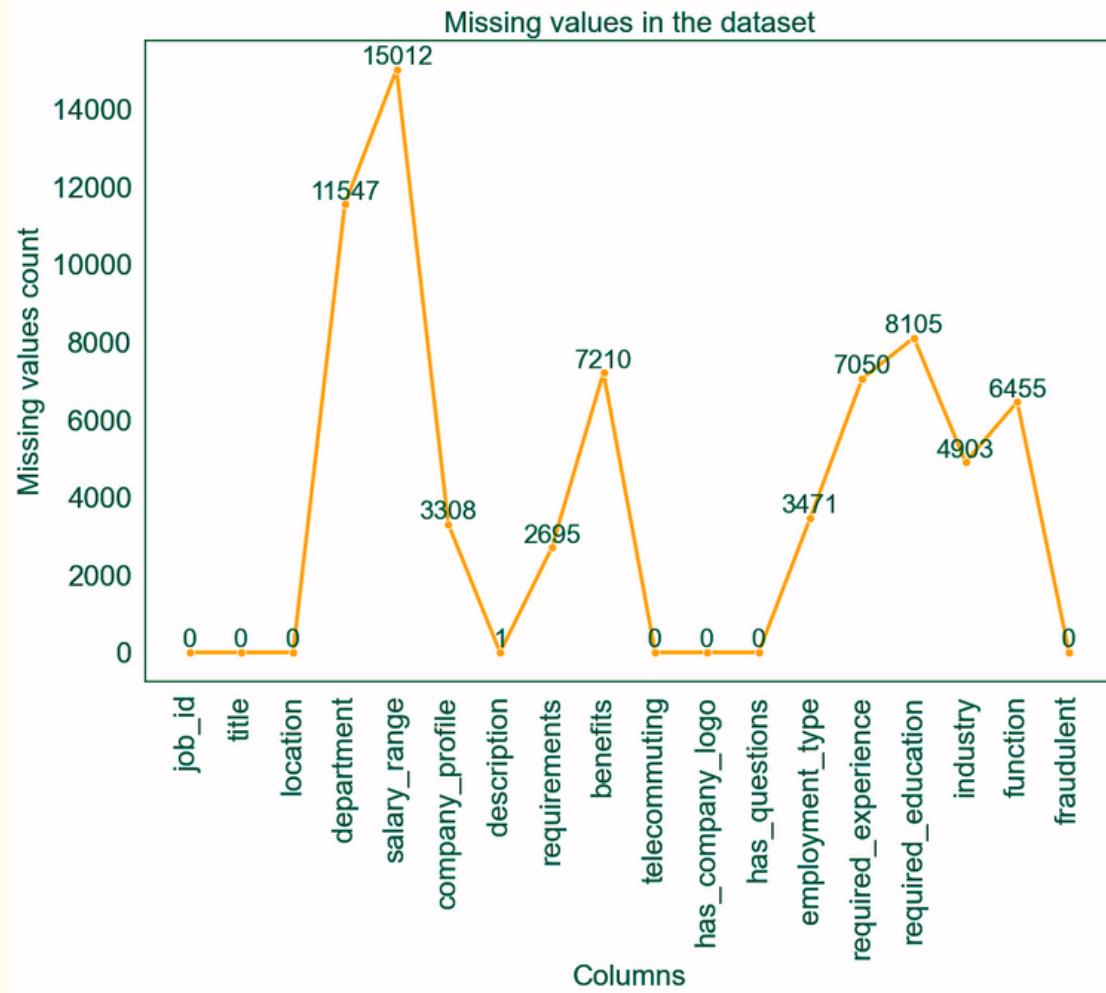


Figure 2: Jobs countries



Number of job postings in different Countries

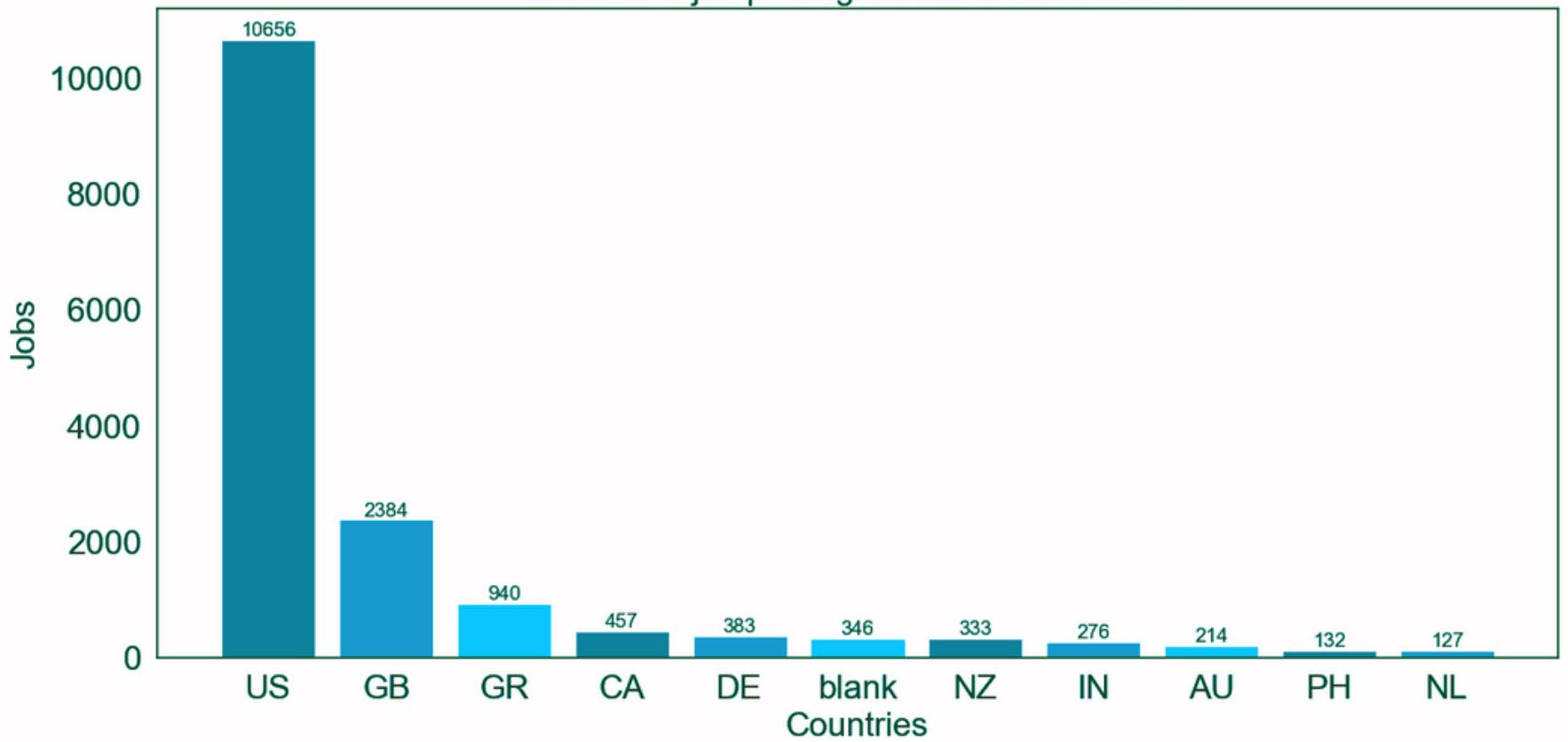
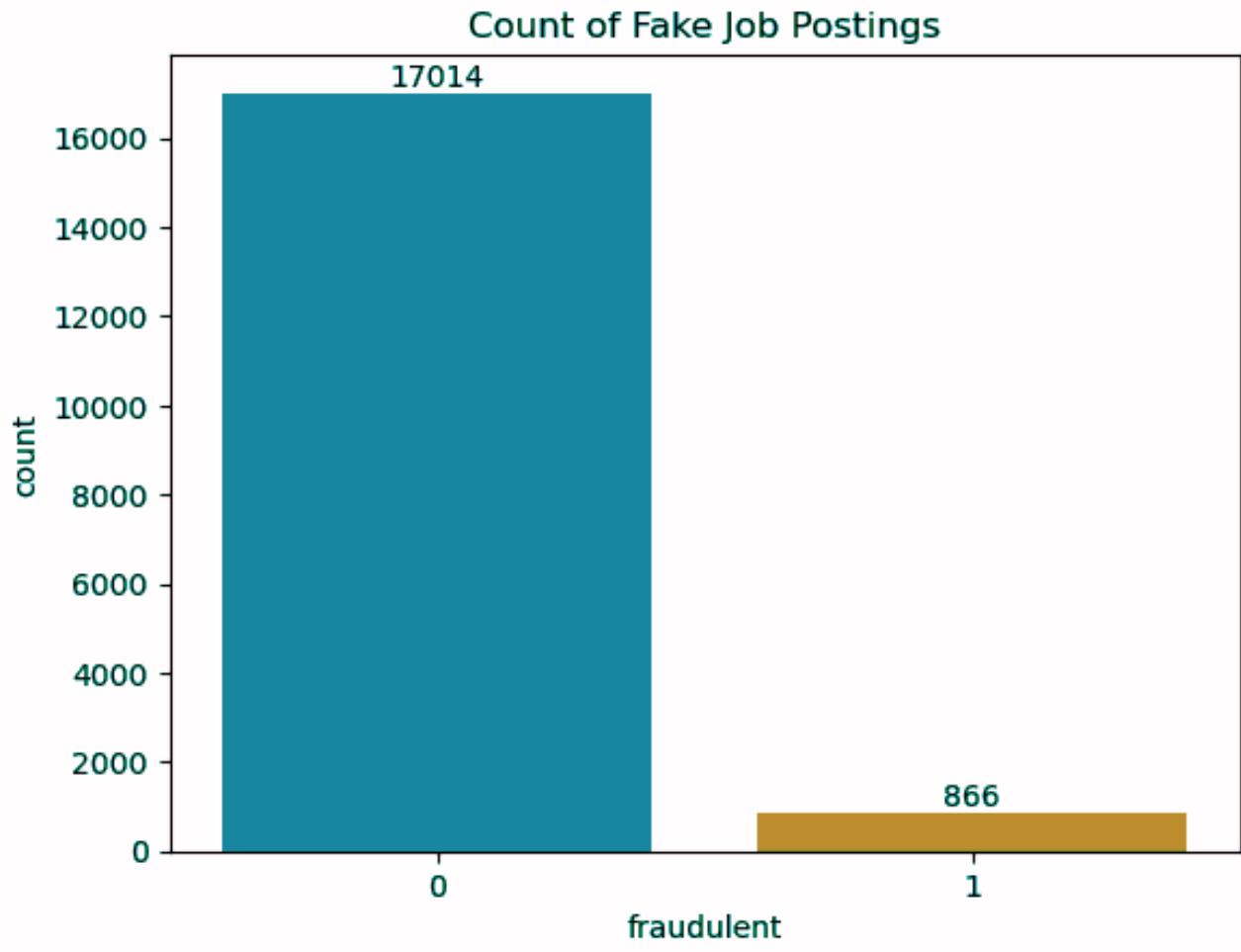


Figure 3: Fraudulent count



1

Treat Missing
values



2

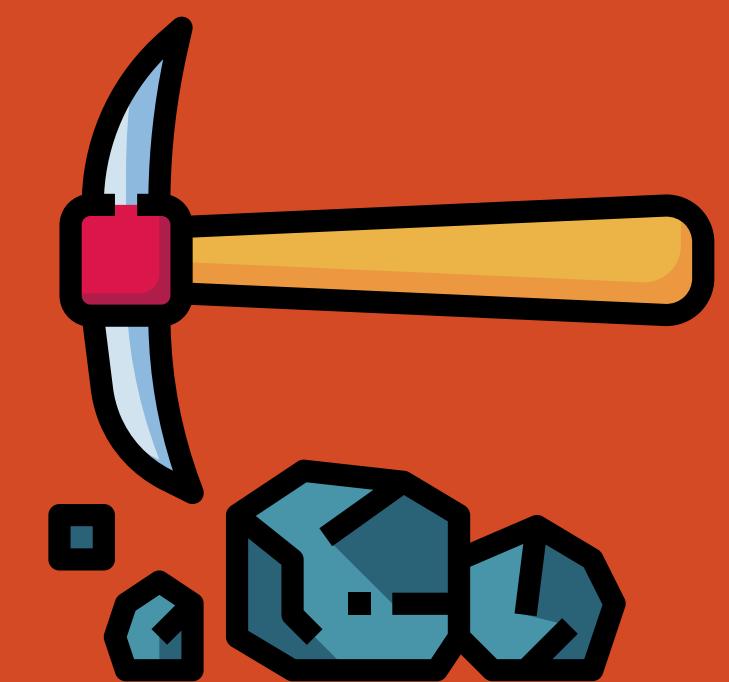
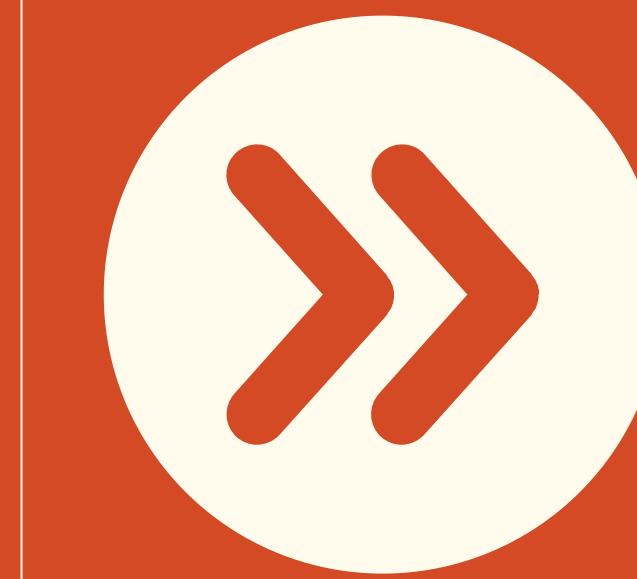
Encoding
categorical Data



3

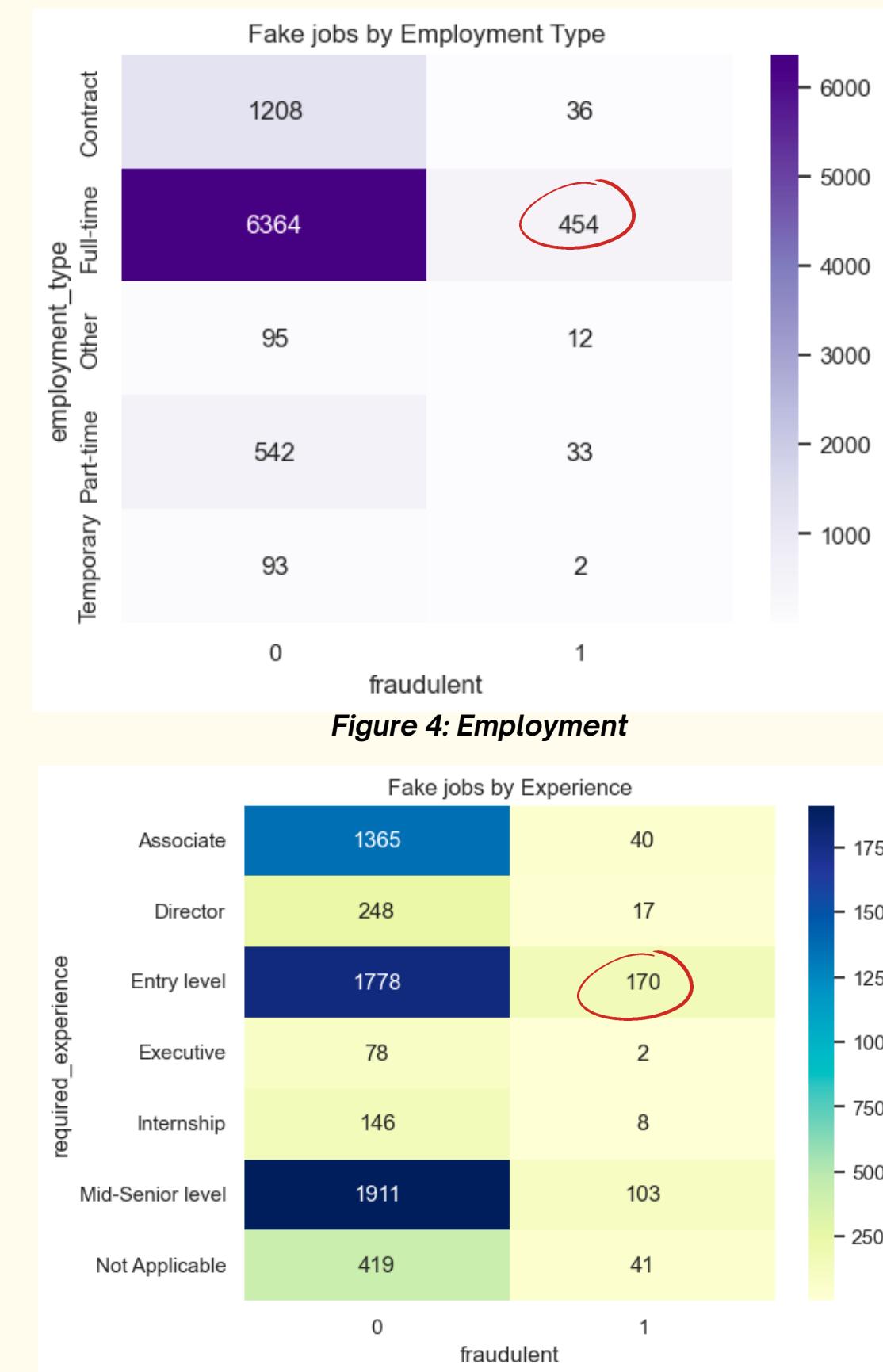
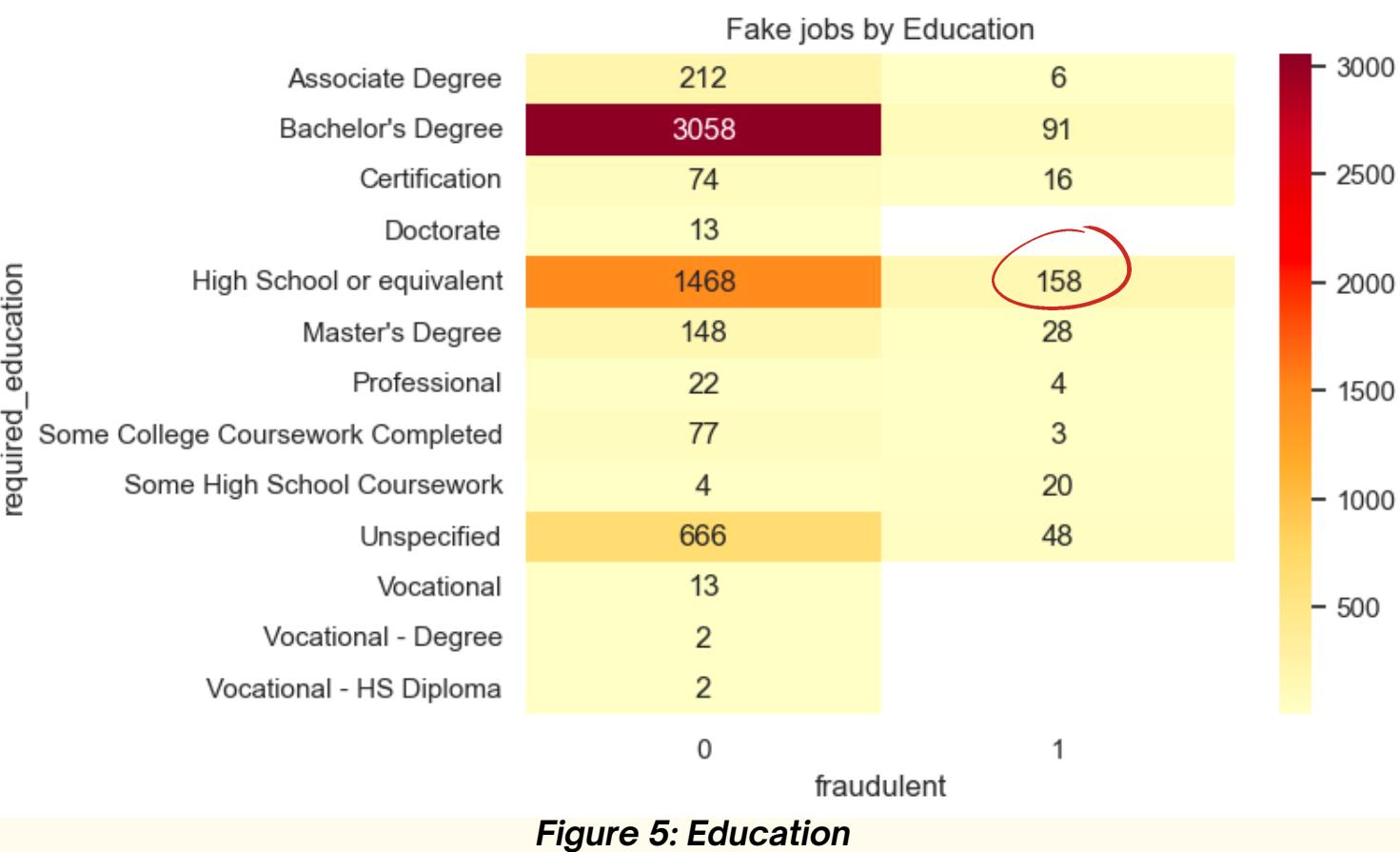
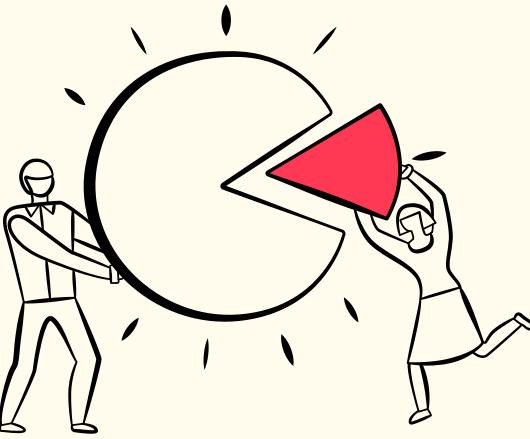
Data Pre- Processing

Text Mining
Techniques



Data ready to go!

Exploratory Data Analysis



US State/City - Statistics

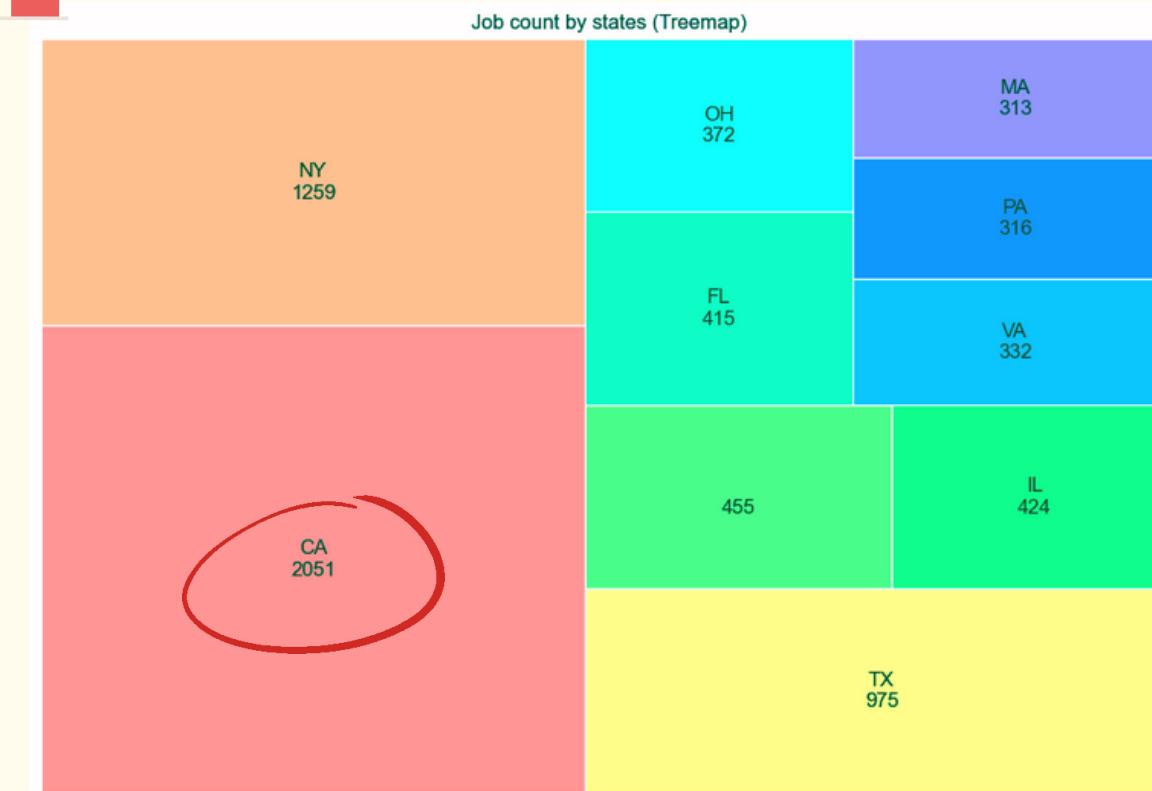


Figure 7 : Job count by States

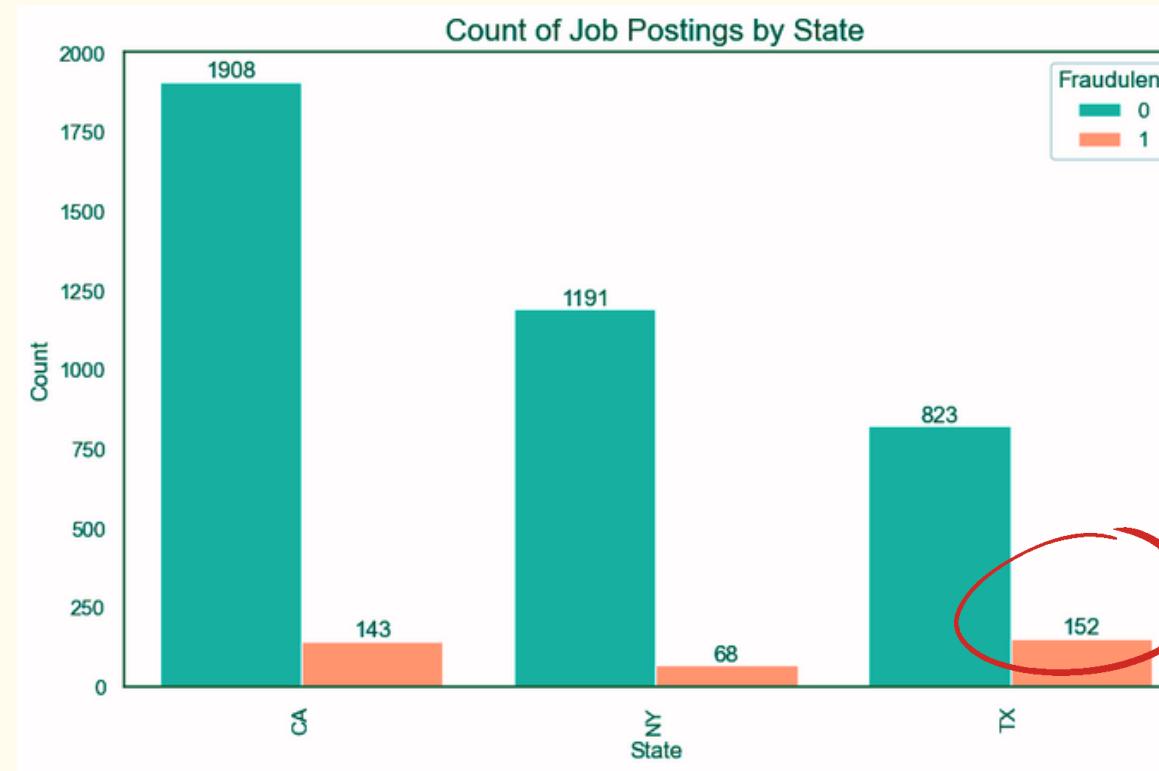


Figure 8: Top 3 States with Fake Jobs

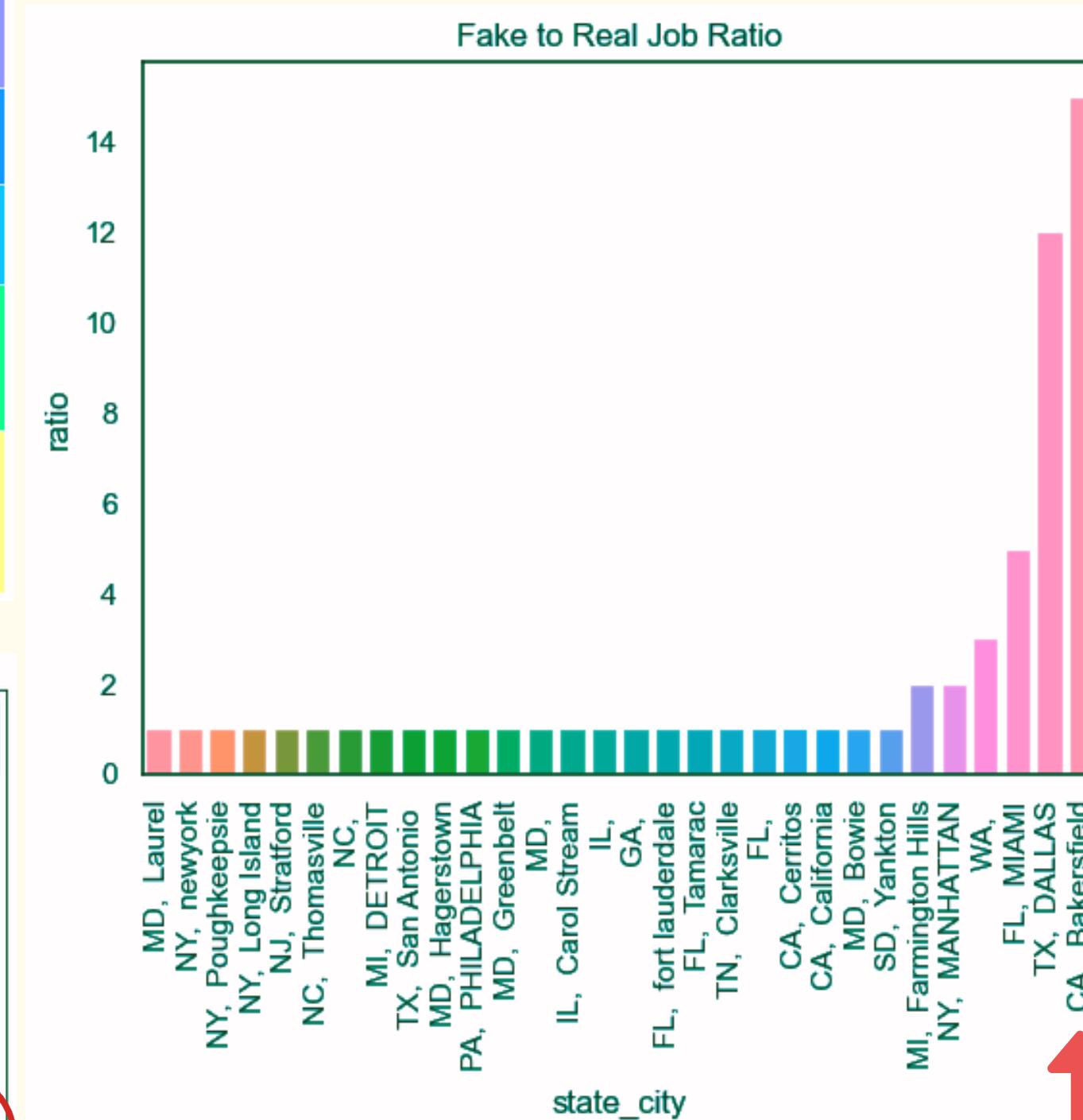


Figure 9: Fake to Real Jobs Ratio - City/State

WORD CLOUD

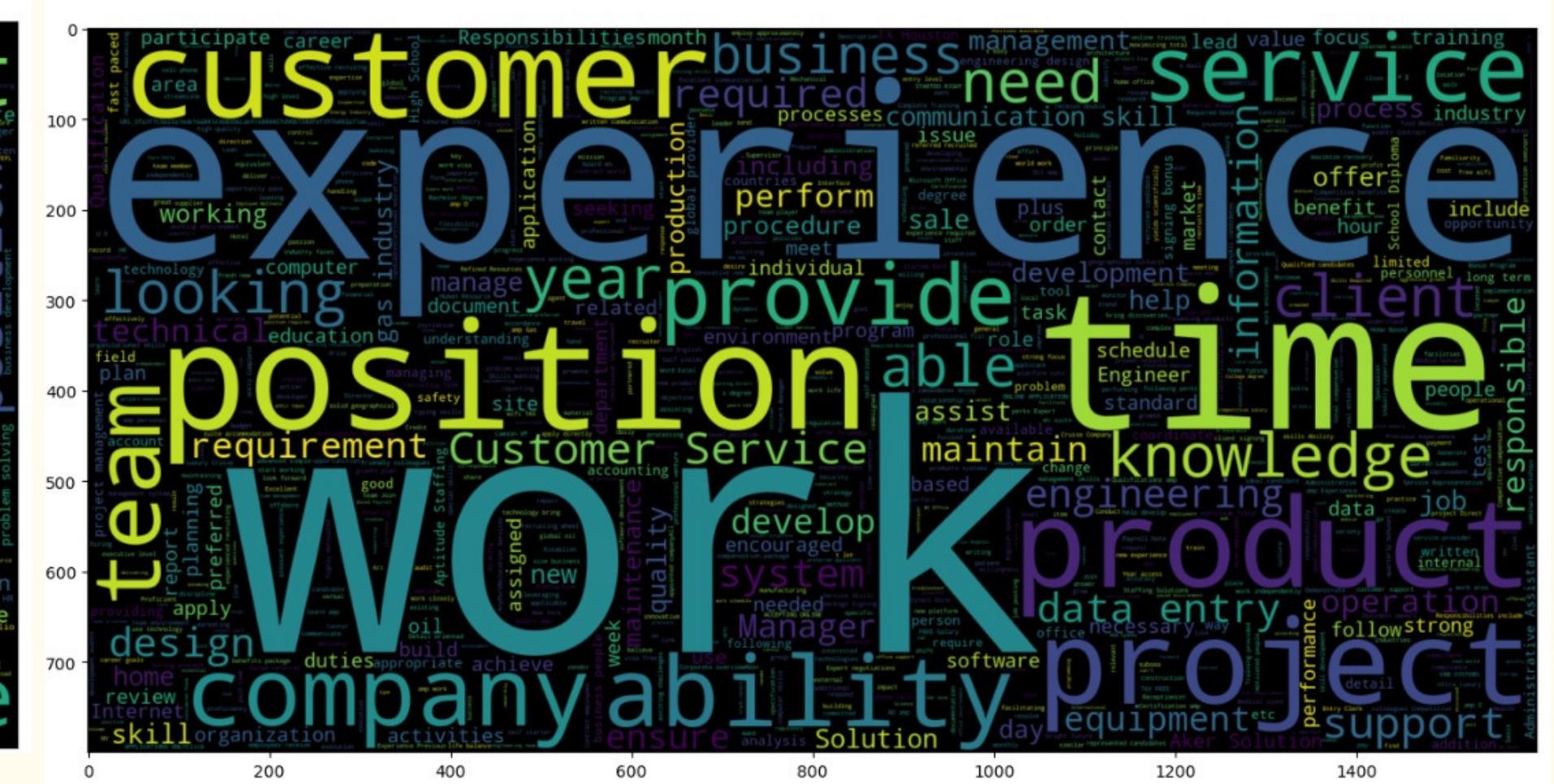


WORD CLOUD

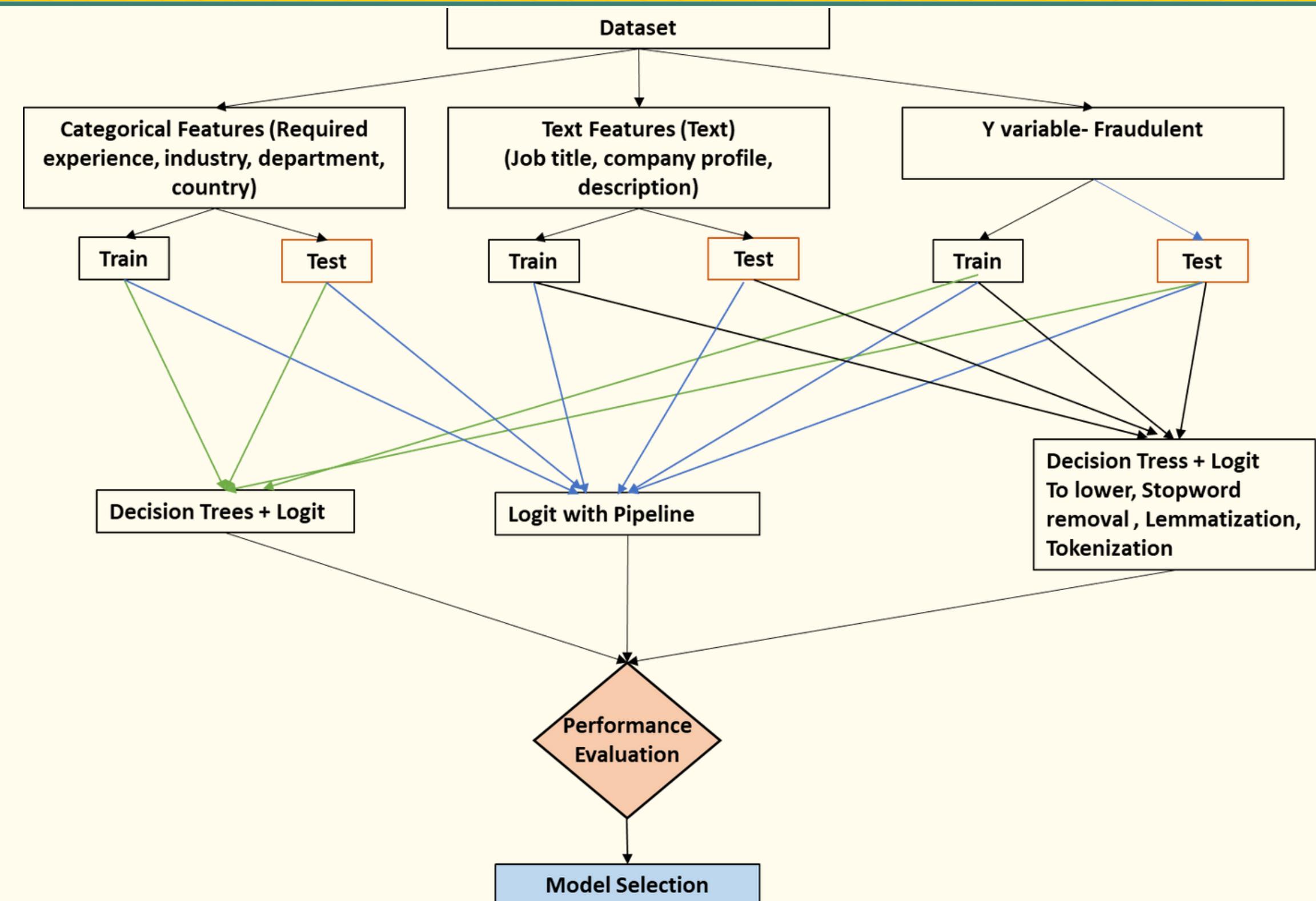
Actual Jobs



Fake Jobs



Model Implementation



Unleashing the power of data with advanced modeling



Different data mining techniques used:

Decision Trees
Booster Trees

**Regression
Models**

LOGIT

Text Mining

KEY WORD ANALYSIS



		Prediction	
Actual	0	1	
0	5081	23	
1	206	54	

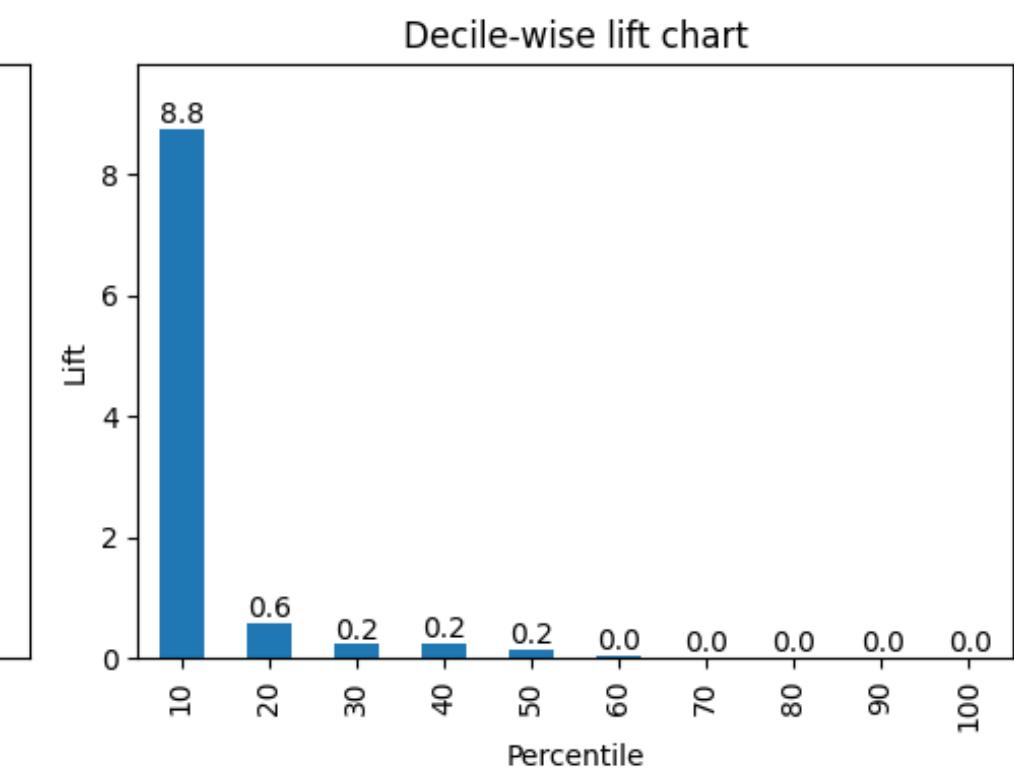
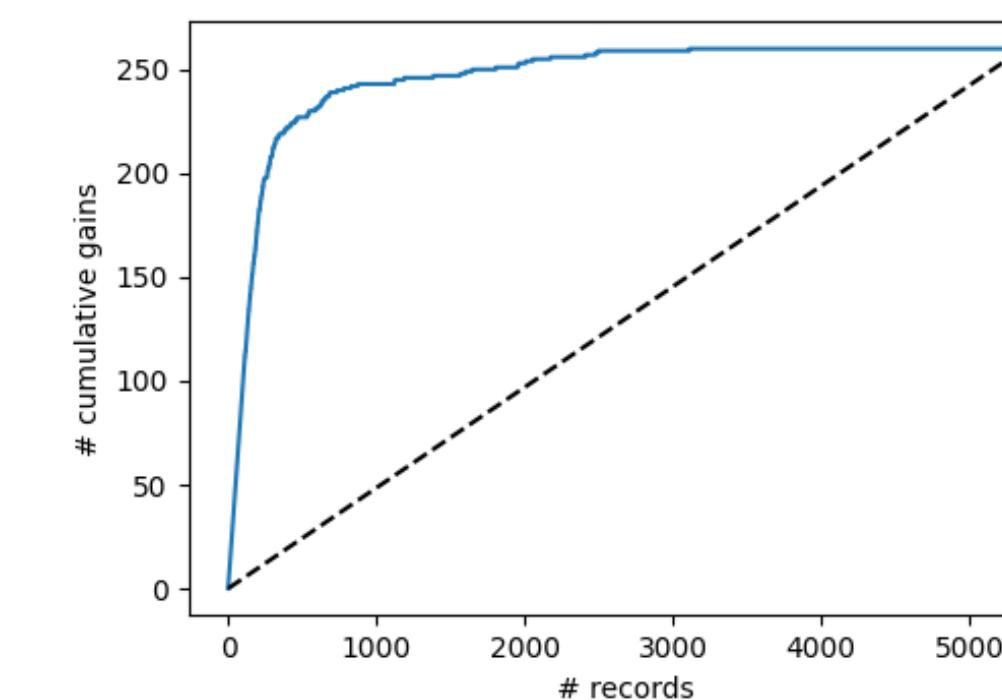
Accuracy: **95.73%**

		Prediction	
Actual	0	1	
0	5100	4	
1	92	168	

Accuracy: **98.21%**

		Prediction	
Actual	0	1	
0	5082	22	
1	64	196	

Accuracy: **98.40%**



Project Learnings

Importance of data cleaning & effective preprocessing

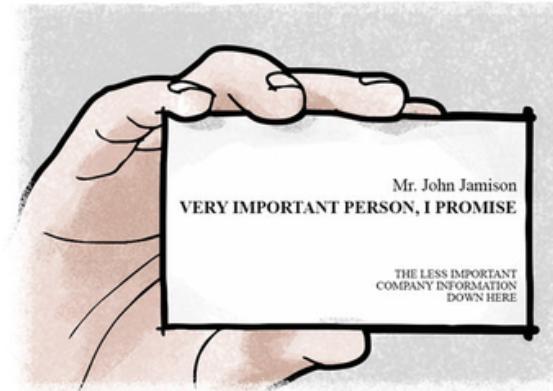
- handling inconsistencies
- handling missing values
- investigating outliers

Improving data mining skills for fraud detection and prevention

- Use of effective text mining
- Machine learning for fraud detection
- Keeping abreast of new developments in data mining and fraud detection techniques.

Fraudulent job posting indicators and red flags

XYZ Company



Vague Job title, job description and company description



Email addresses looking like
[redacted]@my.stjohns.edu

*Too Good
To Be True!*



Moving Forward

We learned a lot with this project and it was great fun too!

Exploring other areas like:

Oversampling

BERT (Bidirectional Encoder Representations from Transformers)

LDA (Latent Dirichlet Allocation).....

