

Project

Gender Prediction and analysis

Presented by - Pinki Raghuwanshi

DATAPLAY

1. Project overview

- Dataset Size
- Data Format
- Data Content

2. Key Features

- Name Analysis
- Predictive modeling

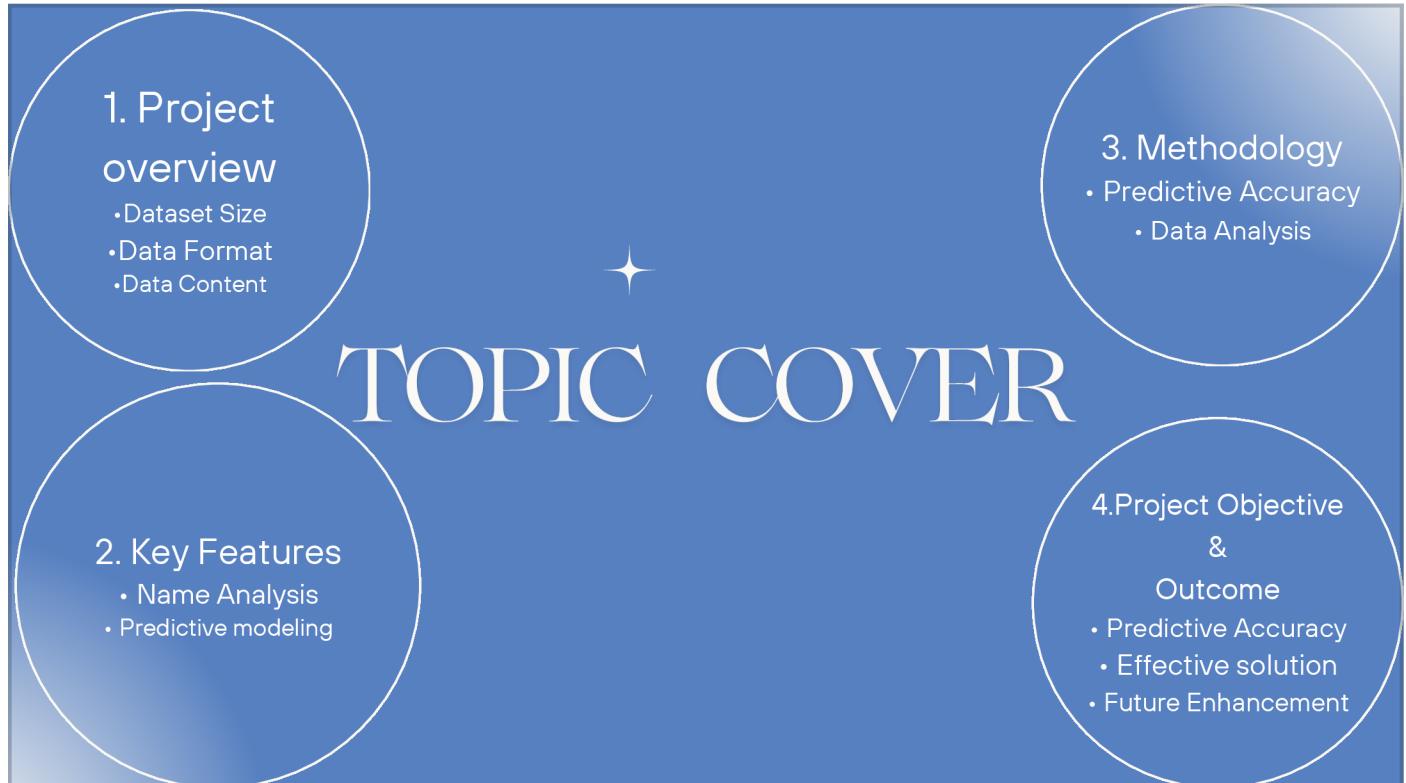
3. Methodology

- Predictive Accuracy
 - Data Analysis

4. Project Objective & Outcome

- Predictive Accuracy
- Effective solution
- Future Enhancement

TOPIC COVER



Project Overview

This project utilizes a comprehensive dataset of 3,000 names, each labeled as Male or Female. The dataset, compiled in an Excel sheet, features a diverse range of names with varying last-letter structures. By leveraging Excel's analytical capabilities, we aim to develop a predictive model that accurately matches names with their corresponding gender labels, demonstrating the practical application of data analysis and predictive modeling using accessible tools.



DATAPLAY

Dataset overview

	Name	Gender	LastLetter
1	Ashutosh	Male	h
3	Meghamala	Female	a
4	Sahib	Male	b
5	Pragya	Female	a
6	Kranti	Female	i
7	Tulika	Female	a
8	Aarushi	Female	i
9	Abhicandra	Male	a
10	Pratigya	Female	a
11	Devak	Male	k
12	Kashiprasad	Male	d
13	Madhavi	Female	i
14	Charusila	Female	a
15	Chithayu	Male	u
16	Manmayi	Female	i
17	Mahajabeen	Female	n
18	Krishnakumar	Male	r
19	Kailas	Male	s
20	Nidhyathi	Female	i
21	Nainika	Female	a
22	Madirakshi	Female	i
23	Adil	Male	l
24	Devadyumna	Male	a
25	Harijatha	Female	a
26	Rupal	Female	l
27	Eka	Female	a
28	Ilampirai	Female	i

Analysis

Add a subheading *Type of Prediction:
Classification (Binary Classification:
Male/Female)

Methodology:* Last Letter Analysis of Names

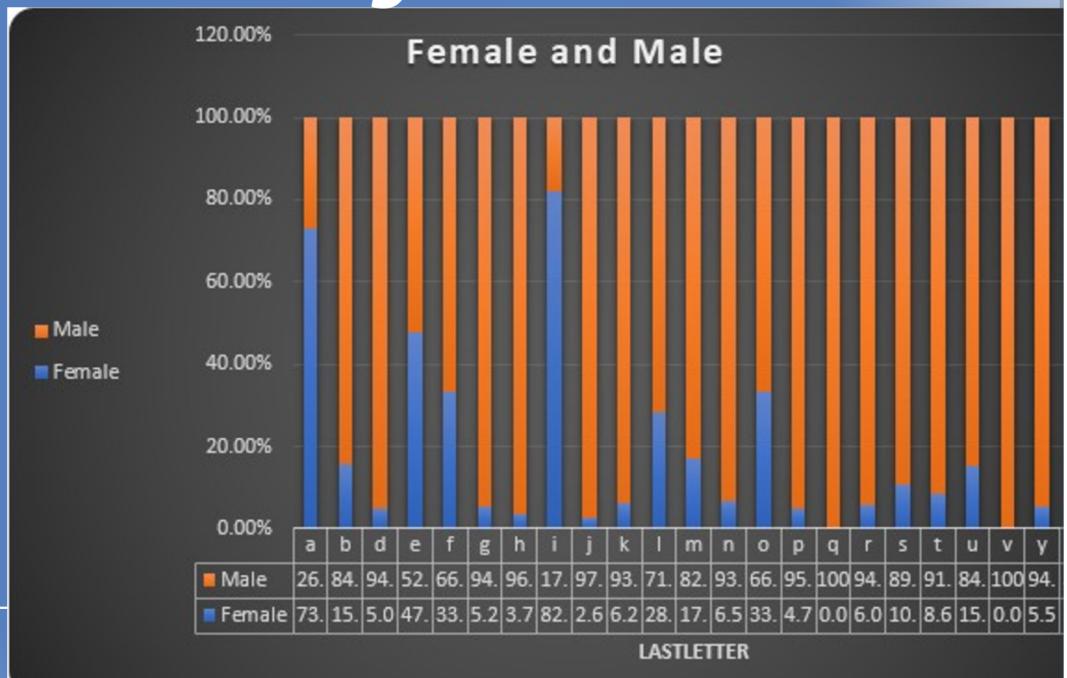
Description:* This project uses the last letter of a name to predict the gender of an individual. The analysis is based on the frequency of last letters in male and female names.

Count of Name LastLetter	Gender		Grand Total
LastLetter	Female	Male	
a	73.15%	26.85%	100.00%
b	15.79%	84.21%	100.00%
d	5.08%	94.92%	100.00%
e	47.62%	52.38%	100.00%
f	33.33%	66.67%	100.00%
g	5.26%	94.74%	100.00%
h	3.72%	96.28%	100.00%
i	82.04%	17.96%	100.00%
j	2.63%	97.37%	100.00%
k	6.25%	93.75%	100.00%
l	28.26%	71.74%	100.00%
m	17.19%	82.81%	100.00%
n	6.54%	93.46%	100.00%
o	33.33%	66.67%	100.00%
p	4.76%	95.24%	100.00%
q	0.00%	100.00%	100.00%
r	6.00%	94.00%	100.00%
s	10.81%	89.19%	100.00%
t	8.65%	91.35%	100.00%
u	15.49%	84.51%	100.00%

Analysis

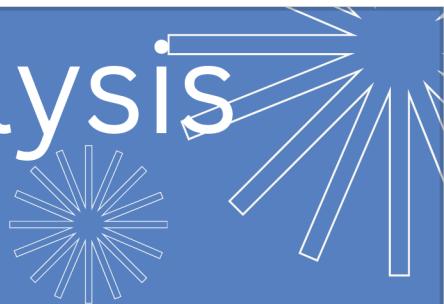
Certain last letters are more commonly associated with male or female names.

- The prediction model achieves [insert accuracy percentage] accuracy in predicting gender based on last letter analysis



Predictive Analysis

Prediction Accuracy



1.1. Male prediction accuracy: 93.46%

2. (for last letter 'N')

2. Female prediction accuracy: 82.04%

(for last letter 'I')

	Name	LastLetter	Male Chances	Female Chances	Gender Prediction
1	Zaraaya	a	26.85%	73.15%	Female
2	Nirvaan	n	93.46%	6.54%	Male
3	Vihaana	a	26.85%	73.15%	Female
4	Saanvi	i	17.96%	82.04%	Female



Prediction Accuracy

Prediction Accuracy

1. Correct Female Predictions: 390

2. Correct Male Predictions: 357

Total Correct Predictions: 747 (390 +357)

4. Total Incorrect Predictions: 153

(29 +124)

Overall Performance

1. Overall Accuracy: 82.78% (747/900)

2. Overall Error Rate: 17.22% (153/900)

Count of CORRECT/INCORRECT	GENDER prediction			
		Female	Male	Grand Total
Female		390	29	419
Male			124	357
Grand Total		514	386	900

precession= TP/ (TP+TN)
recall= TP/(TP+FN)
f1 score= 2(precession x recall / precession+recall)

precession for female : 0.758754864
recall for female: 0.930787589

precession for male : 0.924870466
recall for male: 0.742203742

F1 score female: 0.836012862
F1 score male: 0.823529412

RESULT

Summary of Findings

1. *Last Letter Analysis*: Analyzed the last letter of names to predict gender.
2. *Prediction Accuracy*: Achieved an overall accuracy of 82.78% in predicting gender.
3. *Precision and Recall*: Calculated precision and recall for both female and male predictions.
 - Female: Precision (0.7587), Recall (0.9308), F1 Score (0.8360)
 - Male: Precision (0.9249), Recall (0.7422), F1 Score (0.8235)
4. *Correct and Incorrect Predictions*: Identified 390 correct female predictions, 357 correct male predictions, 29 incorrect female predictions, and 124 incorrect male predictions.

Conclusion

This analysis demonstrates the effectiveness of using last letter analysis to predict gender. The results show a high degree of accuracy, with room for improvement in recall for male predictions.



