

# Pima Indian Diabetes

Group 1 | Prof Itauma | Northwood University



## Problem

Can we build a predictive model to accurately predict whether or not the patients in the dataset have diabetes or not.

## Hypothesis

Patient BMI, pregnancies, Glucose level, DiabetesPedigreeFunction decide whether she is diabetic or not

## Exploratory Data Analysis

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
8	183	64	0	0	23.3	0.672	32	1
5	116	74	0	0	25.6	0.201	30	0
10	115	0	0	0	35.3	0.134	29	0
8	125	96	0	0	0	0.232	54	1
4	110	92	0	0	37.6	0.191	30	0
10	168	74	0	0	38	0.537	34	1
10	139	80	0	0	27.1	1.441	57	0
7	100	0	0	0	30	0.484	32	1
7	107	74	0	0	29.6	0.254	31	1
8	99	84	0	0	35.4	0.388	50	0
7	196	90	0	0	39.8	0.451	41	1
7	147	76	0	0	39.4	0.257	43	1
5	117	92	0	0	34.1	0.337	38	0
6	92	92	0	0	19.9	0.188	28	0
11	138	76	0	0	33.2	0.42	35	0
7	133	84	0	0	40.2	0.696	37	0

## Variables / Research

### Independent variables

- Patient BMI, pregnancies, Glucose level, DiabetesPedigreeFunction

### Dependent variable

- Outcome

## Introduction

- This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage. It's the seventh leading cause of death.1 Healthy People 2030 focuses on reducing diabetes cases, complications, and deaths.

## Model Building

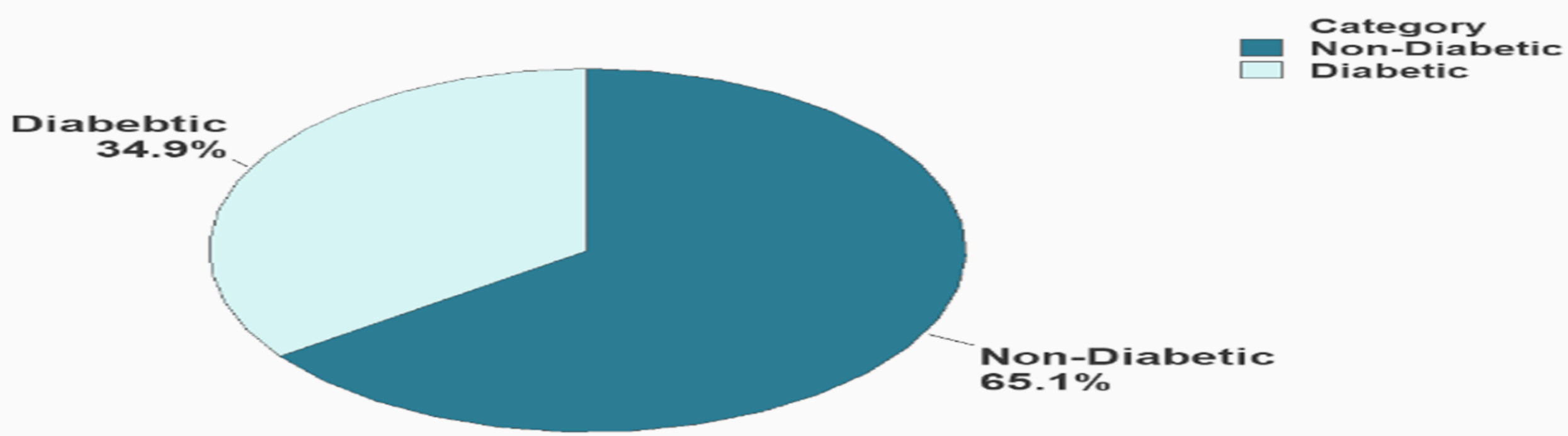
### Odds Ratios for Continuous Predictors

	Odds Ratio	95% CI
Pregnancies	1.1525	(1.0929, 1.2154)
Glucose	1.0344	(1.0276, 1.0412)
BMI	1.0812	(1.0524, 1.1108)
DiabetesPedigreeFunction	2.4628	(1.3904, 4.3623)

### Model Summary

Deviance	Deviance				Area Under
R-Sq	R-Sq(adj)	AIC	AICc	BIC	ROC Curve
27.00%	26.60%	735.20	735.28	758.42	0.8390

Patient Diabetic or Non-Diabetic



## References

## Analysis using different graphs

