Assignment2:

Machine Learning Algorithms to predict Diabetes

- Download the Diabetes Dataset from UCI Machine Learning Repository by clicking on the following link:
  https://archive.ics.uci.edu/ml/machine-learning-databases/diabetes/

- The dataset consists of the following fields:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

Tasks

1. Apply two supervised and two unsupervised Machine Learning algorithms to predict Diabetes using Spark.
2. Use ggplot ( ) to do visual representations of your analysis.