Diabetes Data Story Presentation

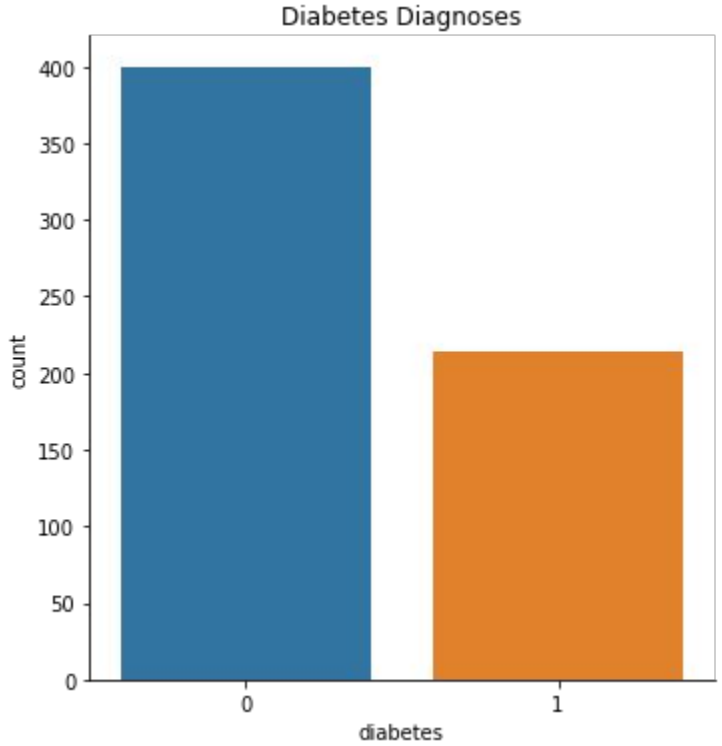
By Sara Maxwell

About the Dataset

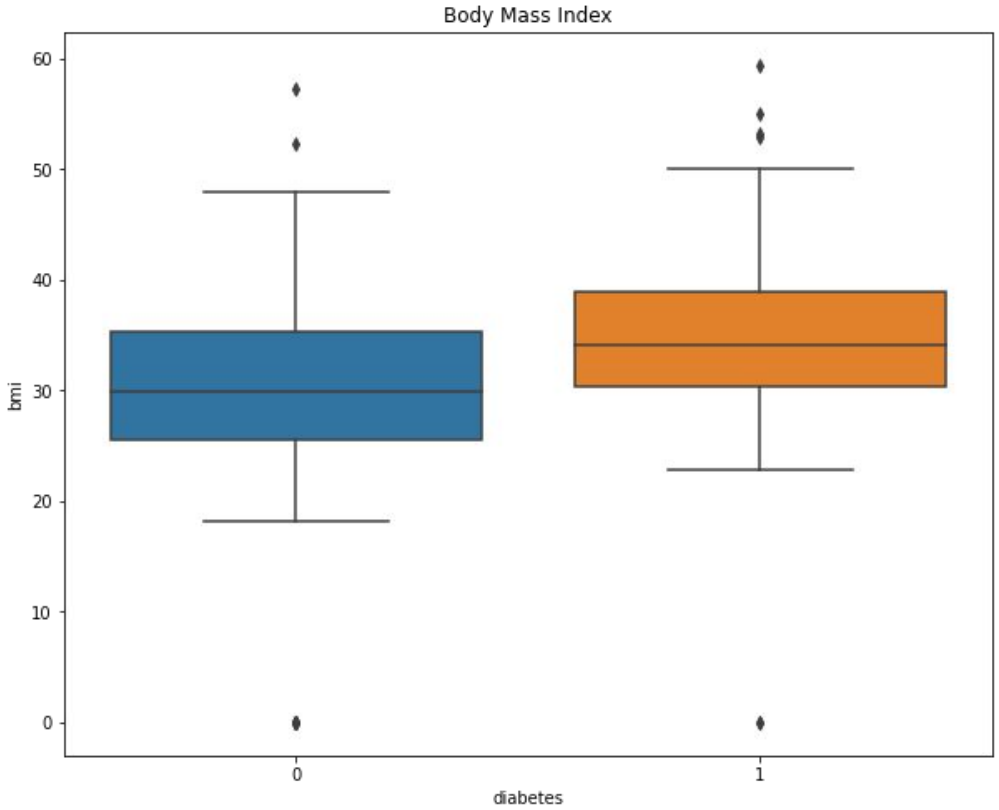
- Binary: Predicts onset of diabetes given certain medical details
- 614 observations
- 8 input variables: number of times pregnant, glucose concentration, blood pressure, skin fold thickness, serum insulin, body mass index, diabetes pedigree, and age
- 1 output variable

Questions

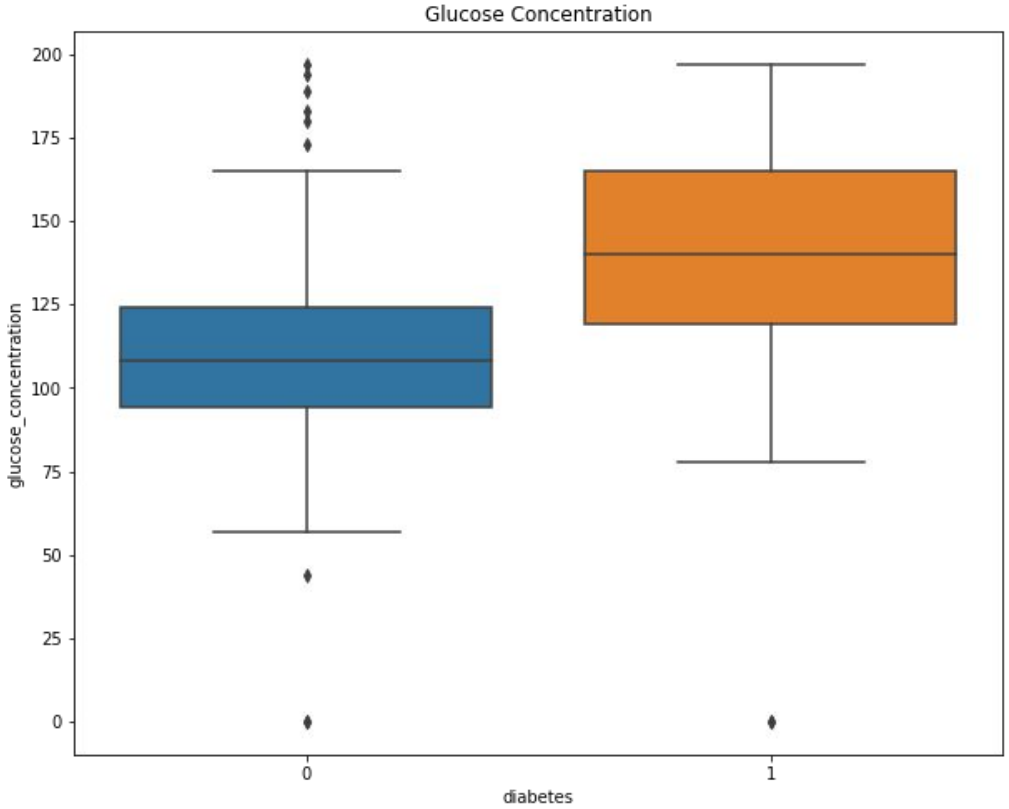
1. Which medical details more strongly predict a positive diabetes diagnosis?
2. Do patients with higher BMIs have higher glucose concentrations?
3. Do patients with higher BMIs have higher serum insulin?
4. Do patients with higher glucose concentrations have higher serum insulin?



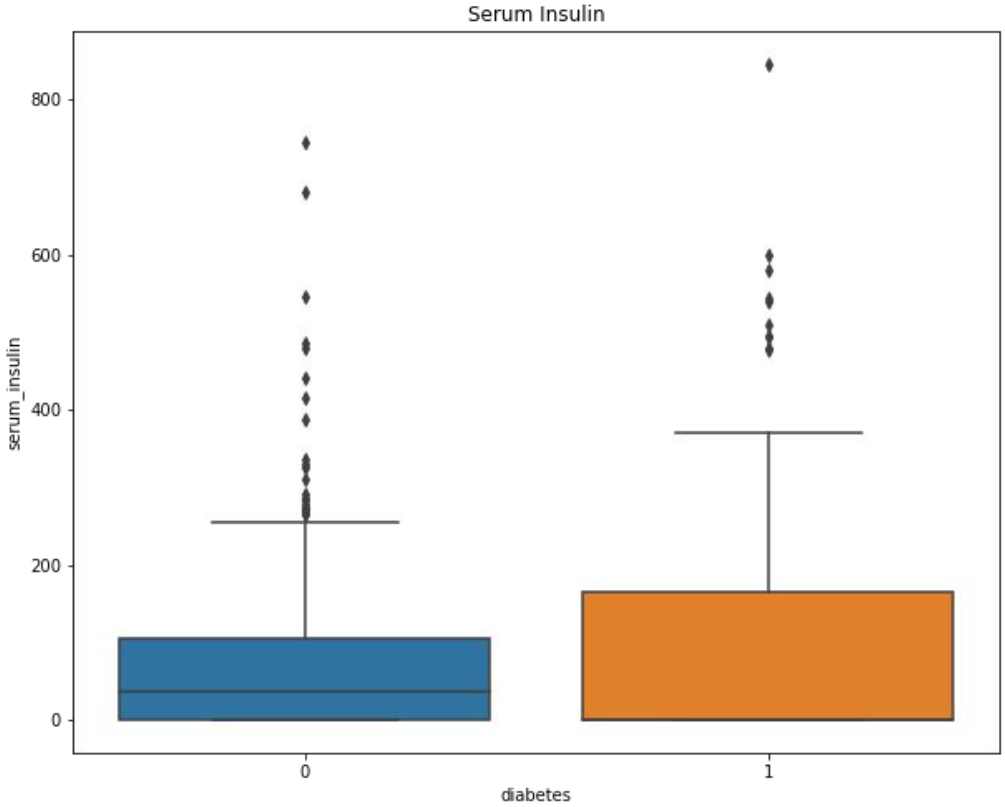
Comparing the number of negative diagnoses of diabetes (0) with positive diagnoses (1)



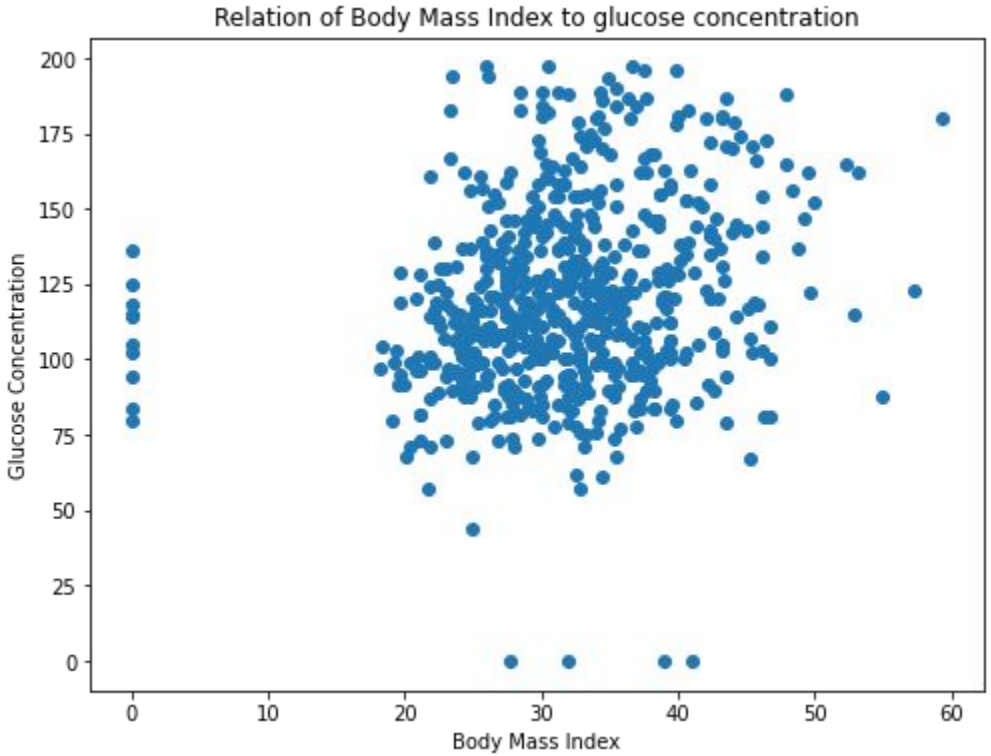
Body mass indices in patients with a negative (0) and a positive (1) diagnosis of diabetes



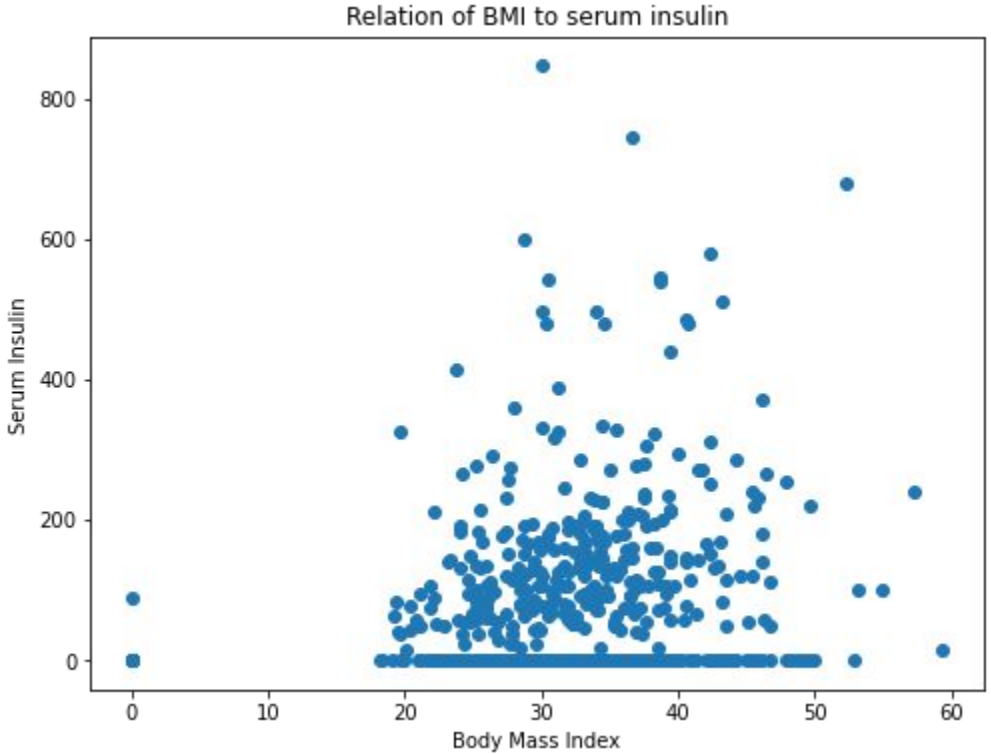
Blood glucose concentration in patients with a negative (0) and a positive (1) diagnosis of diabetes.



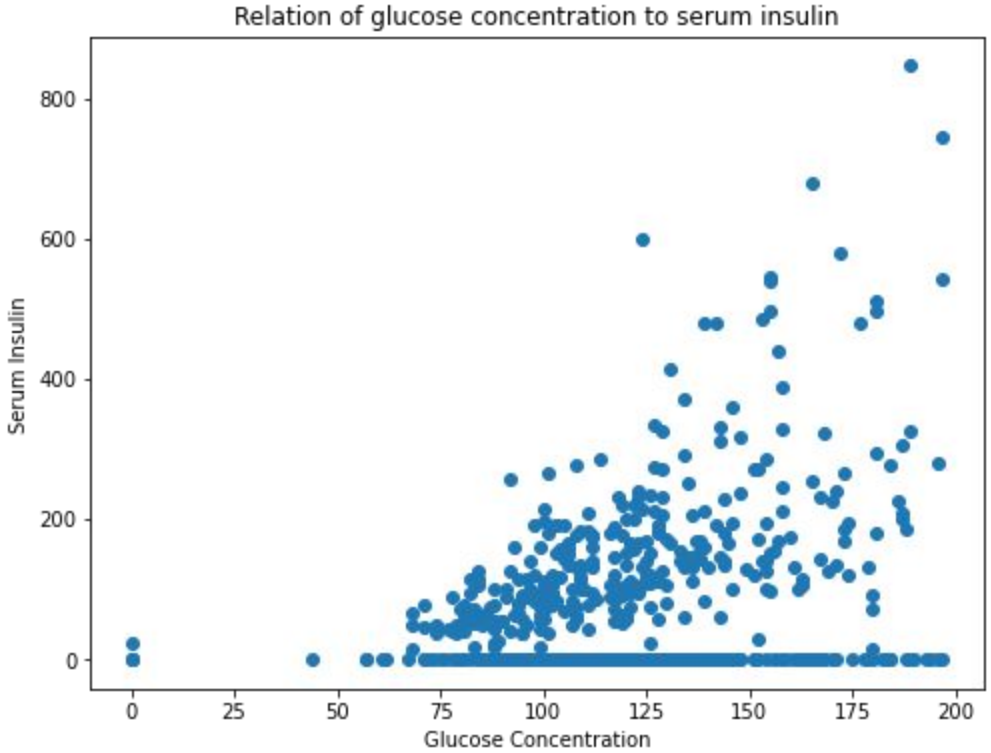
Serum insulin in patients with a negative (0) and a positive (1) diagnosis of diabetes.



Relation of body mass index to glucose concentration. There isn’t a strong correlation between body mass index and glucose concentration.



Relation of body mass index to serum insulin. There isn’t a strong correlation between body mass index and serum insulin.



Relation of glucose concentration to serum insulin. There is a correlation between glucose concentration and serum insulin.

Conclusions

- Factors most likely to predict a positive diabetes diagnosis are serum insulin and glucose concentration
- Body mass index is not correlated with glucose concentration or serum insulin
- Glucose concentration is correlated with serum insulin