Introduction to Data Science
Winter 2018

Midterm Exam

- You have 45 minutes to complete the exam
- The exam is closed book, closed notes, closed computer, closed calculator, except one hand-written 8.5 × 11 crib sheet of your own creation and the official study guide provided with the exam.
- Mark your answers **on the exam itself**. We will **not** grade answers written on scratch paper. If you run out of room, continue on the back of the page.

Name

Student id number

Email

All the work on this exam is my own
(Please sign)

Maternal smoking and pregnacies

Printed below are the first three rows of a table. The column labels should be self explanatory.

```
baby = Table().read_table("baby.csv")
baby.show(3)
```

Birth Weight	Gestational Days	Maternal Age	Maternal Height	Maternal Pregnancy Weight	Maternal Smoker
120	284	27	62	100	False
113	282	33	64	135	False
128	279	28	64	115	True

... (1171 rows omitted)

In the following you may assume that the statements **from data science import** $\star$ and **import numpy as np** have been executed.

Problem 1 (10 points)

Write Python code that computes the birth weight of the heaviest baby.

Problem 2 (10 points)

Write Python code that computes the average birth weight of babies of 33 year old mothers.

Problem 3 (10 points)

Write Python code that computes the average birth weight of babies of 33 year old smokers.

Problem 4 (10 points)

Write Python code that draws a bar chart showing the mean birth weights for babies of smokers and babies of non-smokers.

Problem 5 (20 points)

Write a function

`mean_birth_weight(query_age)`

that returns the average birth weight of babies with Maternal Age = query_age if there are such babies in the data set, and "None" otherwise.

Hint: Using **`Table.where`** returns a table with 0 rows if there are no rows matching the predicate you supply.

Problem 6 (10 points)

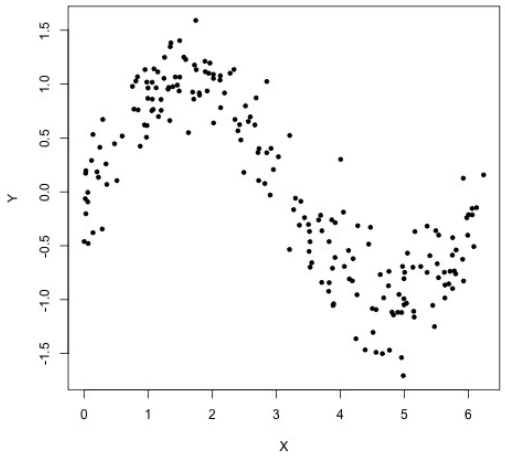
The average birth weight for babies of smokers is 114oz, while the average birth weight for babies of non-smokers is 123oz. Based solely on this information, can you conclude that smoking causes low birth weight? Briefly justify your answer.

Problem 7 (10 points)

In 1954 the Public Health Service conducted a study to test the effectiveness of Jonas Salk's polio vaccine. The incidence of polio among children who were vaccinated was 28 per 100,000, while the incidence among unvaccinated children was 71 per 100,000. The Public Health Service concluded that the vaccine does indeed reduce the incidence of polio. What is the crucial aspect of the study that justifies this conclusion?

Problem 8 (5 points)

In the figure below, draw the conditional mean of Y, given X.



Problem 9 (5 points)

In the figure below, draw the condition mean of Y, given X.