Homework 1: Data Visualization

INSTRUCTOR: DANIEL L. PIMENTEL-ALARCÓN DUE 01/31/2019

In this homework you will obtain some insights from data through visualization. To this end you will use the titanic dataset (`titanic_data.csv`), containing the following information about 887 passengers on board the Titanic ship: 1) whether they survived or not (1 = survived, 0 = deceased), 2) passenger class, 3) gender (0 = male, 1 = female), 4) age, 5) number of siblings/spouses aboard, 6) number of parents/children aboard, and 7) fare:

	Passenger 1	Passenger 2	Passenger 3	$\cdots$	Passenger 887
Survived	**0**	**1**	**1**	$\cdots$	**0**
Passenger Class	3	1	3	$\cdots$	3
Gender	0	1	1	$\cdots$	0
Age	22	38	26	$\cdots$	32
Siblings/Spouses	1	1	0	$\cdots$	0
Parents/Children	0	0	0	$\cdots$	0
Fare	7.25	71.2833	7.925	$\cdots$	7.75

(a) Plot a histogram detailing the *age* distribution of the dataset. What do you conclude from this histogram?

(b) Show a violin plot detailing the *age* distribution, separated by survivors and deceased. What do you conclude from this plot?

(c) Show a box plot summarizing the *fare* distribution, separated by survivors and deceased. What do you conclude from this plot?

(d) Show a scatter plot of *gender* vs. *survival*. What do you conclude from this plot?

(e) Plot a histogram showing the distribution of survivors as a function of gender. What do you conclude from this histogram?

(f) Plot a histogram showing the distribution of gender as a function of survival. What do you conclude from this histogram?