## CSCI183 HW2 Report

Codes for all other steps are in the ipynb file.

3. there are 9 numerical features, and 1 categorical feature (ocean proximity), assuming no target specified

4.

```
I find high corr between the following pairs:
corr(longitude, latitude) = -0.924664 but I think this is because of the geographic location of california
corr(total_rooms, total_bedrooms) = 0.930380
corr(total_rooms, population) = 0.857126
corr(total_rooms, households) = 0.918484
corr(total_bedrooms, population) = 0.877747
corr(total_bedrooms, households) = 0.979728
corr(population, households) = 0.907222
```

6. Based on the graphs in step 5, identify features that have a linear relationship with the target variable.

```
(total_rooms, total_bedrooms)
(total_rooms, population)
(total_rooms, households)
(total_bedrooms, population)
(total_bedrooms, households)
(population, households)
(total_bedrooms, total_rooms)
(population, total_rooms)
(population, total_rooms)
(population, total_rooms)
(population, total_bedrooms)
(households, total_bedrooms)
(households, population)
```

Feature	Target	Slope/ Coefficient	Intercept	Graph
total_bedrooms	total_rooms	4.82242124	46.89997601810228	total_bedrooms vs total_rooms  40000 -

population	total_rooms	1.65119792	282.01885268768046	population vs total_rooms	_
				60000 -	
				50000 -	
				30000	
				40000 -	
				s a	
				tota - 10000 -	
				tota for the state of the state	
				20000 -	
				10000 -	
				0-	
				0 5000 10000 15000 20000 25000 30000 35000	0
				population	
households	total_rooms	5.24097265	17.689280170347047	households vs total_rooms	
households	total_rooms	5.24097265	17.689280170347047	households vs total_rooms	
households	total_rooms	5.24097265	17.689280170347047		
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 -	,
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 -	,
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 - 25000 - 8, 20000 - 15000 - 10000 -	9
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 - 25000 - 80, 20000 - 151, 20000 - 152, 20000 -	,
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 - 25000 - E 2 15000 - 10000 - 5000 - 0 -	•
households	total_rooms	5.24097265	17.689280170347047	40000 - 35000 - 30000 - 25000 - 8, 20000 - 15000 - 10000 -	9

population	total_bedrooms	0.3235017	75.69470115245974	population vs total_bedrooms
				12000
				10000 -
				8000 -
				D. 6000 -
				1 ped 2000 - 600
				g 4000 -
				2000 -
				0-1
				0 5000 10000 15000 20000 25000 30000 35000
				population
households	total_bedrooms	1.06881901	2.9213477383798363	
Tiouseriolus	total_bediooilis	1.00001901	2.9213477303790303	households vs total_bedrooms
				6000 -
				6000 -
				5000 -
				5000 -
				5000 -
				5000 -
				1 total 2 and 2 and 3 and 4 and 5 an
				5000 -
				1 total 2 and 2 and 3 and 4 and 5 an
				5000 - SE 4000 - TE 3000 - 1000 - 1000 -
				5000 -
				5000 - SE 4000 - TE 3000 - 1000 - 1000 -

households	population	2.68719566	83.11588362697034	households vs population	
				35000 -	•
				30000 -	•
				25000 -	
				20000 - population 15000 -	
				od 15000 -	
				10000 -	
				5000 -	
				0 -	
					0 1000 2000 3000 4000 5000 6000 households