Stat 1350, Regression Problems, Fall 2014   Name ___KEY___

graphs on attached page

Instructions: For each of the data sets below, determine the following:
  i.   Draw a scatterplot of the data
  ii.   Is the graph showing a linear or non-linear relationship
  iii.   For linear graphs, find the least-squares regression line and the correlation coefficient.
  iv.   What proportion of the data is explained by the linear relationship?
  v.   Predict an appropriate measurement.
  vi.   Give an example of a measurement you would not want to use with this model.
  vii.   How strong is the relationship between the two variables?
  viii.   Are there any obvious outliers?

1. No tortilla chip aficionado likes soggy chips, so it is important to find characteristics of the production process that produce chips with an appealing texture. The following data on x=frying time (sec) and y=moisture content (%) appeared in an article on the subject.

x	5	10	15	20	25	30	45	60
y	16.3	9.7	8.1	4.2	3.4	2.9	1.9	1.3

ii. nonlinear

2. The following data on y=glucose concentration (g/L) and x=fermentation time (days) for a particular blend of malt liquor was read from a scatterplot in an article.

x	1	2	3	4	5	6	7	8
y	74	54	52	51	52	53	58	71

ii. nonlinear

3. The data below compared cricket chirps with temperature.

X	Y
20	88.6
16	71.6
19.8	93.3
18.4	84.3
17.1	80.6
15.5	75.2
14.7	69.7
17.1	82
15.4	69.4
16.2	83.3
15	79.6
17.2	82.6
16	80.6
17	83.5
14.4	76.3

ii. linear
iii. $y = 3.291x + 25.23$; $r = .835$
iv. $r^2 = .697 \sim 69.7\%$
v. $x = 22 \Rightarrow y = 3.291(22) + 25.23 = 97.63$
vi. $x = 30$ or $x = 5$ (answers may vary)
vii. pretty strong
viii. no

4.

Hours Spent Studying	Math SAT Score
4	390
9	580
10	650
14	730
4	410
7	530
12	600
22	790
1	350
3	400
8	590
11	640
5	450
6	520
10	690
11	690
16	770
13	700
13	730
10	640

5.

House Price in $1000s (Y)	Square Feet (X)
245	1400
312	1600
279	1700
308	1875
199	1100
219	1550
405	2350
324	2450
319	1425
255	1700

4.

ii. linear

iii. $y = 25.33x + 353.2$; $r = .9336$

iv. $r^2 = .8716$ ≈ 87.2%

v. $y = 25.33(15) + 353.2 = 733$

vi. $x = 35$

vii. very strong

viii. yes $x=22$ look like one correlation goes even higher if this point removed.

---

5. ii linear

iii. $y = .10976x + 98.248$; $r = .76$

iv. $r^2 = .5808$ ≈ 58%

v. $y = .10976(2000) + 98.248 = $317,790$
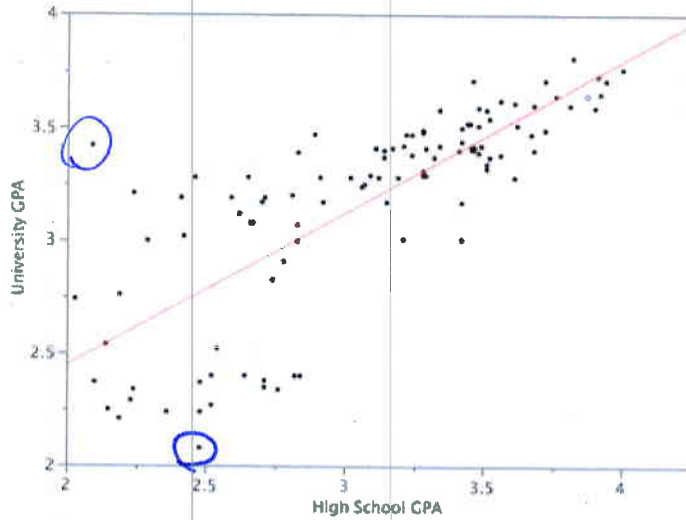
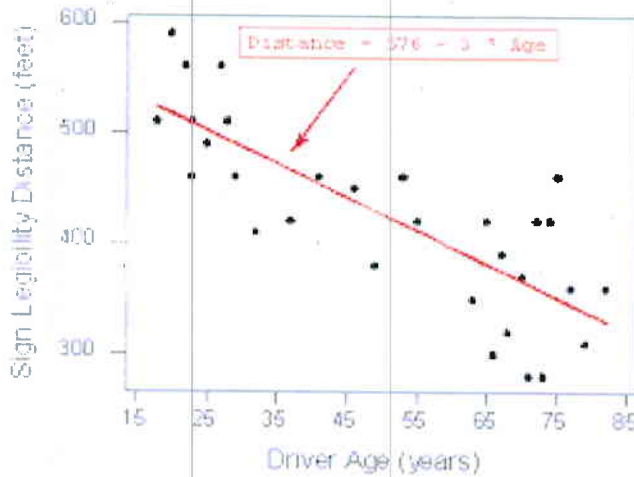vi. $x = 4000$ or $x = 500$

vii. moderate

viii. no

For each of the scatterplots below, determine if the best-fit equation is linear or non-linear. If linear, if the correlation positive, negative or approximately zero? Are there any outliers?
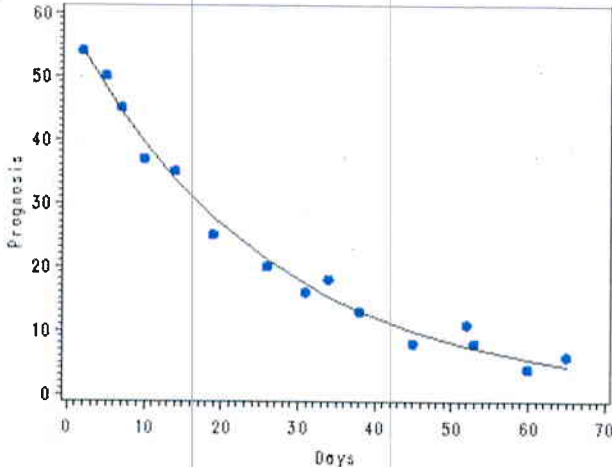
6.

linear, though there
is more variation for
smaller values.
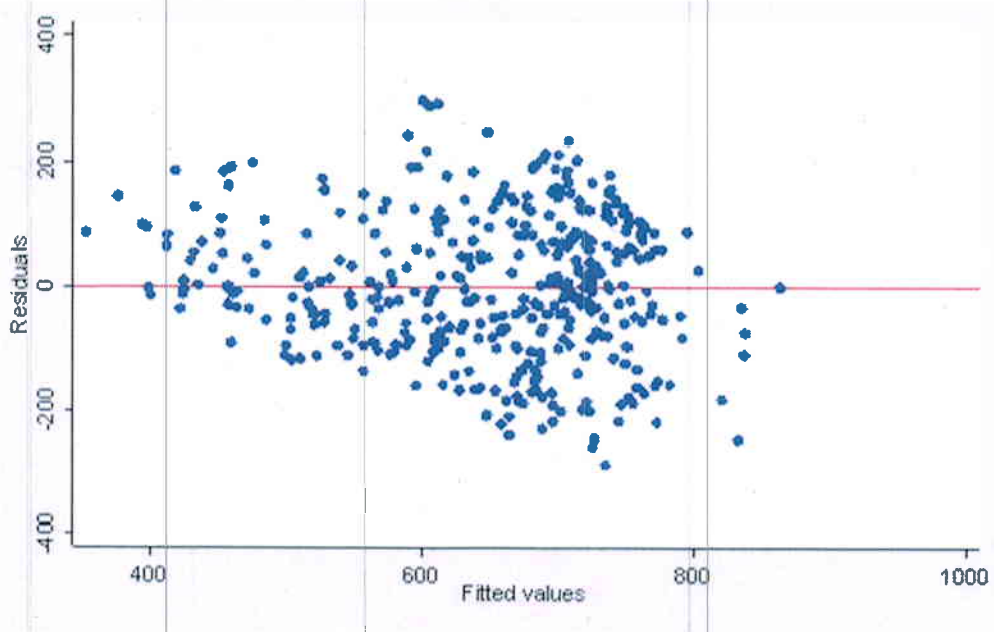positive correlation
possible outliers circled

7.

linear
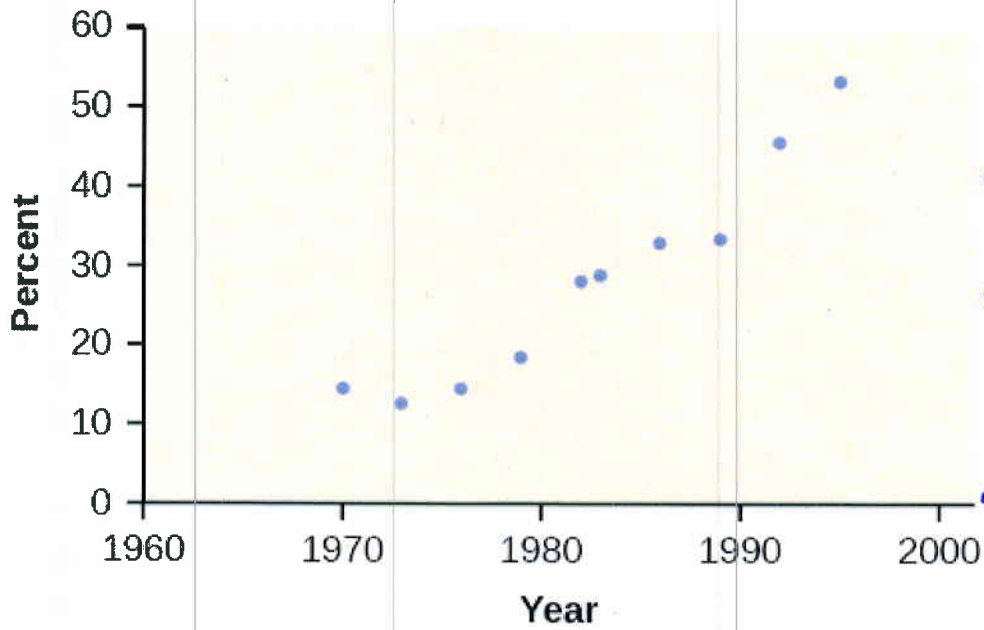negative correlation
no obvious outliers

8.

nonlinear

9. .

no correlation

10. .

Student Ethnic Minority Percentage

possibily linear w/ 1970 pt. as outlier or could be nonlinear correlation (if linear) is positive

1.

2.

3.

4.

5.