## Ch 12 Homework

1. The life expectancy for an individual born in the United States in cer

| Year of Birth | Life Expectancy |
|---------------|-----------------|
| 1930 | 59.7 |
| 1940 | 62.9 |
| 1950 | 70.2 |
| 1965 | 69.7 |
| 1973 | 71.4 |
| 1982 | 74.5 |
| 1987 | 75 |
| 1992 | 75.7 |
| 2010 | 78.7 |

birth is independent variable
life expectany is dependent variable

Correlation is: 0.96126116 Life expectancy increases

H0:there isn't a linear correlation between year of birth and life expectancy
H1:there is a linear correlation between year of birth and life expectancy
Pvalue<5% we rejected H0
conclusion: there is a linear correlation between year of birth and life expectancy

0.2275*1850-377.2431=43.6319

**Parameter estimates:**

| Parameter | Estimate | Std. Err. | Alternative | DF | T-Stat | P-value |
|-----------|----------|-----------|-------------|----|--------|---------|
| Intercept | -377.2431 | 48.570088 | $\neq 0$ | 7 | -7.766984 | 0.0001 |
| Slope | 0.22747972 | 0.024654334 | $\neq 0$ | 7 | 9.2267639 | <0.0001 |

Possible lurking variables: Medical advancements, nutrition
These can significantly influence life expectancy
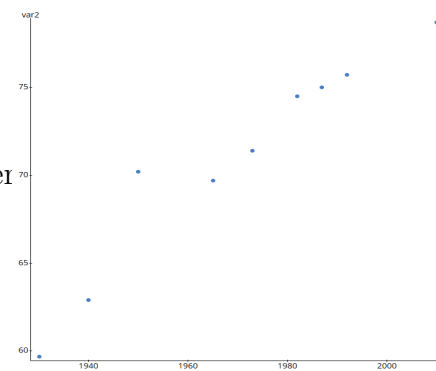
a) Which variable should be the independent variable and which should be the dependent variable?
b) Sketch a scatterplot. Do points form a straight line?
c) Find the correlation coefficient $r$. What does it suggest about the slope? About the strength of the association?
d) Test the claim that there is a linear correlation. Write an appropriate conclusion.
e) Find the estimated life expectancy for an individual born in 1850.
f) Find a lurking/confounding variable if possible. Explain why it may change the response variable.

2. The table below lists weight of a diamond and the price associated to that weight of 6 dimonds.

weight is independent variable
Priceis dependent variable

| 3 | 0.4 | 0.5 | 0.5 | 1.0 | 0.7 |
|---|-----|-----|-----|-----|-----|
| 0 | 1151 | 1343 | 1410 | 5669 | 2277 |

le should be the independent variable and which should be the dependent

terplot. Do points form a straight line? yes its straight

relation coefficient $r$. What does it suggest about the slope? About the strength of the association? r=0.96763687 Suggests that for each 1 carat increase, the price goes up by about $7,177.

d) Test the claim that there is a linear correlation. Write an appropriate conclusion.
e) Not included in the table below is a diamond with a weight of 1.50 carats. Find the best predicted price for this diamond. Is the result close to the current actual price?
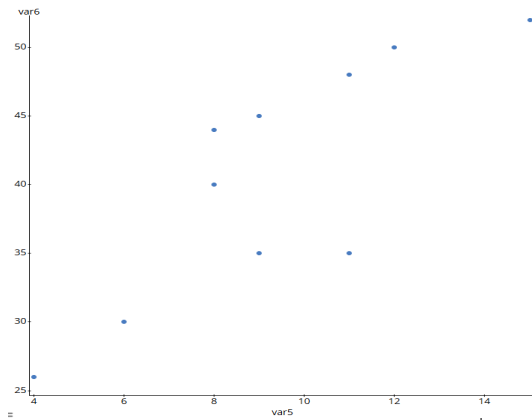
H0:there isn't a linear correlation
H1:there is a linear correlation
Pvalue<5% we rejected H0
there is a linear correlation

7177.0213*1.50-2006.9787=8758,55
it might underestimate the true value

| Parameter | Estimate | Std. Err. | Alternative | DF | T-Stat | P-value |
|-----------|----------|-----------|-------------|----|--------|---------|
| Intercept | -2006.9787 | 571.80498 | $\neq 0$ | 4 | -3.5099008 | 0.0247 |
| Slope | 7177.0213 | 935.83557 | $\neq 0$ | 4 | 7.6691051 | 0.0016 |
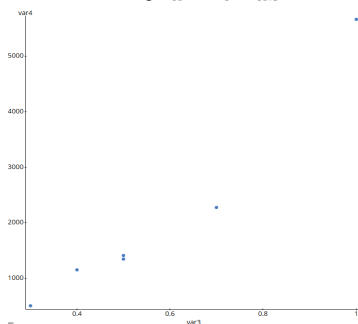
**Parameter estimates:**

| Parameter | Estimate | Std. Err. | Alternative | DF | T-Stat | P-value |
|-----------|----------|-----------|-------------|----|--------|---------|
| Intercept | 19.123723 | 5.5408082 | $\neq 0$ | 8 | 3.4514321 | 0.0087 |
| Slope | 2.2985244 | 0.56757945 | $\neq 0$ | 8 | 4.0496963 | 0.0037 |

tudents study before an exam and the associated scores
s).

| 12 | 4 | 9 | 11 | 8 | 9 | 8 | 15 | 11 |
|----|---|---|----|---|---|---|----|----|
| 50 | 26 | 45 | 48 | 40 | 35 | 44 | 52 | 35 |

a) Which variable should be the independent variable and which should be the dependent variable?  *hours is independent variable*
*Score dependent variable*
b) Sketch a scatterplot. Do points form a straight line?  *yes it's straight*
c) Find the correlation coefficient $r$. What does it suggest about the slope?  About the strength of the association?  *r=0.81983638 for each 1 hours study will up 2.298 points*

*H0:there isn't a linear correlation*
*H1:there is a linear correlation*
d) Test the claim that there is a linear correlation. Write an appropriate conclusion.  *Pvalue<5% we rejected H0*
*there is a linear correlation*
e) Find the score on the exam if a student studies for 16 hours before the exam.
f) Find a lurking/confounding variable if possible. Explain why it may change the response variable.  *2.2985*16+19.1237=55.8997*

*The more time you spend studying, the more knowledge you can acquire you won't lost so many point on the part you studied a lot*

4. The years of working and the associated annual income of college graduates.

| Years of working | 1 | 3 | 4 | 9 | 2 | 8 | 9 | 8 | 15 | 11 |
|------------------|----|----|----|------|------|-----|------|-----|------|-----|
| Annual income | 45k | 50k | 57k | 100k | 120k | 80k | 130k | 90k | 175k | 95k |

where 1k=$1,000 and 10k=$10,000

a) Which variable should be the independent variable and which should be the dependent variable?  *hours is independent variable*
*Score dependent variable*
b) Sketch a scatterplot. Do points form a straight line? *yes it's straight*
c) Find the correlation coefficient $r$. What does it suggest about the slope?  About the strength of the association?  *r=0.73208238 for work  more each 1 year salary will up 6642 dollar*
d) Test the claim that there is a linear correlation. Write an appropriate conclusion.
e) Find the annual income of a college graduate who have been working for 5 years.
f) Find a lurking/confounding variable if possible. Explain why it may change the response variable.
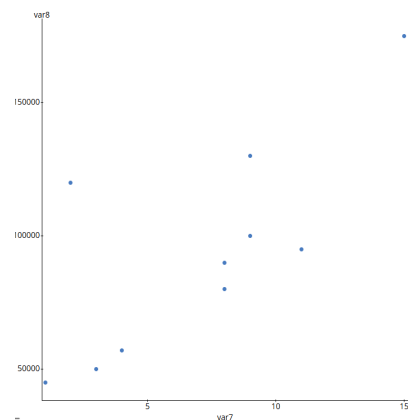
*H0:there isn't a linear correlation*
*H1:there is a linear correlation*
*Pvalue<5% we rejected H0*
*there is a linear correlation*

*6642.0455*5+47705.682=80915.9095*

*Work performance become better*

**Parameter estimates:**

| Parameter | Estimate | Std. Err. | Alternative | DF | T-Stat | P-value |
|-----------|----------|-----------|-------------|----|--------|---------|
| Intercept | 47705.682 | 17832.775 | $\neq 0$ | 8 | 2.6751688 | 0.0281 |
| Slope | 6642.0455 | 2185.1528 | $\neq 0$ | 8 | 3.0396252 | 0.0161 |

Page 2