

Correlation and Regression*

Worksheet 3

1. For a given day, the number of heating degrees is 65 minus the mean temperature in degrees Fahrenheit. If the mean is above 65, the number of heating degrees is 0 for that day. *Heating degree days* is the sum of the heating degrees over a certain period (say a month). They are commonly used in calculations relating to the energy consumption required to heat buildings. We will use linear regression to predict kilowatt hour (kWh) energy use from heating degree days (hdd) (for a particular house in Scotland). Here are the data.

Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
hdd	163	228	343	373	301	238	137	84	38	15	14	34
kWh	593	676	1335	1149	1127	892	538	289	172	131	134	134

- (a) Which is the explanatory variable and which is the response?
 - (b) Display a scatterplot and describe any structure you see in the data.
 - (c) Give the equation of the regression line and place it on a scatterplot.
 - (d) What is the predicted energy use for a month in which the average temperature is 50°F and no day has an average temperature above 65°F?
 - (e) In a cooler climate, the baseload energy or non-weather-dependent consumption is the amount of energy used when none is devoted to heating. Estimate this using the regression line.
 - (f) Find the correlation of these two quantitative variables. What does this value for the correlation tell you?
2. Global warming has many indirect effects on climate. For example, summer monsoon winds in the Arabian Sea bring rain to India needed for agriculture. As the climate warms and winter snow cover in Europe and Asia decreases, the land heats up more rapidly in summer which may increase the strength of the monsoon. The data are snow cover (in millions of square kilometers) and stress (in newtons per square meter).

```
> snowcover<-c(6.6,5.9,6.8,7.7,7.9,7.8,8.1,16.6,18.2,15.2,16.2,17.1,17.3,  
+ 18.1,26.6,27.1,27.5,28.4,28.6,29.6,29.4)  
> windstress<-c(0.125,0.160,0.158,0.155,0.169,0.173,0.196,0.111,0.106,0.  
+ 143,0.153,0.155,0.133,0.130,0.062,0.051,0.068,0.055,0.033,0.029,0.024)
```

- (a) Give the correlation between snow cover and wind stress.
- (b) Give a scatterplot with snow cover as the explanatory variable.

*©2019 Joseph C Watkins

- (c) Find the equation for the corresponding regression line and add it to the scatterplot.
- (d) Use this to predict wind stress when snow cover is 12 million square kilometers.
- (e) Give the regression line using wind stress as the explanatory variable and add it to the scatterplot.
- (f) Take the predicted value for wind stress in part (d) and use this to predict snow cover. Is the value 12 million square kilometers?
- (g) What does your findings in part (f) tell you?