Story Name: Smoking and Cancer

 Abstract: Government statisticians in England conducted a study of the relationship between smoking and lung cancer. The data concern 25 occupational groups and are condensed from data on thousands of individual men. The explanatory variable is the number of cigarettes smoked per day by men in each occupation relative to the number smoked by all men of the same age. This smoking ratio is 100 if men in an occupation are exactly average in their smoking, it is below 100 if they smoke less than average, and above 100 if they smoke more than average. The response variable is the standardized mortality ratio for deaths from lung cancer. It is also measured relative to the entire population of men of the same ages as those studied, and is greater or less than 100 when there are more or fewer deaths from lung cancer than would be expected based on the experience of all English men.  A scatterplot of the data shows a moderately strong linear association, with a correlation coefficient of 0.716. Residuals from a regression of mortality on smoking are randomly scattered with no outliers or influential observations.  Reference: Moore, David S., and George P. McCabe (1989). Introduction to the Practice of Statistics. Original source: Occupational Mortality: The Registrar General's Decennial Supplement for England and Wales, 1970-1972, Her Majesty's Stationery Office, London, 1978. Authorization: Description: Data summarizes a study of men in 25 occupational groups in England. Two indices are presented for each occupational group. The smoking index is the ratio of the average number of cigarettes smoked per day by men in the particular occupational group to the average number of cigarettes smoked per day by all men. The mortality index is the ratio of the rate of deaths from lung cancer among men in the particular occupational group to the rate of deaths from lung cancer among all men. Number of cases: 25 Variable Names:

1.Occupational\_Group: Occupational Group  2.Smoking: Smoking index (100 = average)  3.Mortality: Lung cancer mortality index (100 = average)

**The Data**:

|  |  |  |
| --- | --- | --- |
| OccupationalGroup | Smoking | Mortality  |
| Farmersforestersandfisherman | 77 | 84  |
| Minersandquarrymen | 137 | 116  |
| Gascokeandchemicalmakers | 117 | 123  |
| Glassandceramicsmakers | 94 | 128  |
| Furnaceforgefoundryandrollingmillworkers | 116 | 155 |
|  Electricalandelectronicsworkers | 102 | 101 |
|  Engineeringandalliedtrades | 111 | 118  |
| Woodworkers | 93 | 113  |
| Leatherworkers | 88 | 104  |
| Textileworkers | 102 | 88  |
| Clothingworkers | 91 | 104  |
| Food,drink,andtobaccoworkers | 104 | 129  |
| Paperandprintingworkers | 107 | 86  |
| Makersofotherproducts | 112 | 96  |
| Constructionworkers | 113 | 144 |
|  Paintersanddecorators | 110 | 139  |
| Driversofstationaryengines,cranes,etc. | 125 | 113 |
|  Laborersnotincludedelsewhere | 133 | 146  |
| Transportandcommunicationsworkers | 115 | 128  |
| Warehousemenstorekeeperspackersandbottlers | 105 | 115 |
|  Clericalworkers | 87 | 79  |
| Salesworkers | 91 | 85  |
| Servicesportandrecreationworkers | 100 | 120  |
| Administratorsandmanagers | 76 | 60  |
| Professionals,technicalworkers,andartists | 66 | 51  |