Hanford part #1 - Line of Best Fit

Introduction

In an article taken from the Journal of Environmental Health, May-June 1965, Volume 27, Number 6, pages 883- 897, author Robert Fadely explains that the Atomic Energy Plant in Hanford, Washington has been a plutonium production facility since the Second World War. Some of the waste have been stored underground in the same area. Radioactive waste has been seeping into the Columbia River, and eight Oregon counties and the city of Portland have been exposed to radioactive contamination. The table below lists the number of cancer deaths per 100,000 residents for Portland and these counties. The table also includes an index of exposure that measures the proximity of the residents to the contamination. The index is based on the assumption that city or county exposure is directly proportional to river frontage and inversely proportional both to the distance from Hanford, WA site and to the square of the county's or city's average distance from the river.

Formatting the Data

```
In[23]:= index = {2.5, 2.6, 3.4, 1.3, 1.6, 3.8, 11.6, 6.4, 8.3};

In[24]:= deaths = {147, 130, 130, 114, 138, 162, 208, 178, 210};

In[25]:= location = List["Umatilla", "Morrow", "Gilliam", "Sherman",
            "Wasco", "Hood River", "Portland", "Columbia", "Clatsop"];

In[44]:= data = Grid[{location, index, deaths},
           Frame → All, ItemStyle → Directive[FontSize → 12, Bold]]
```

Out[44]=

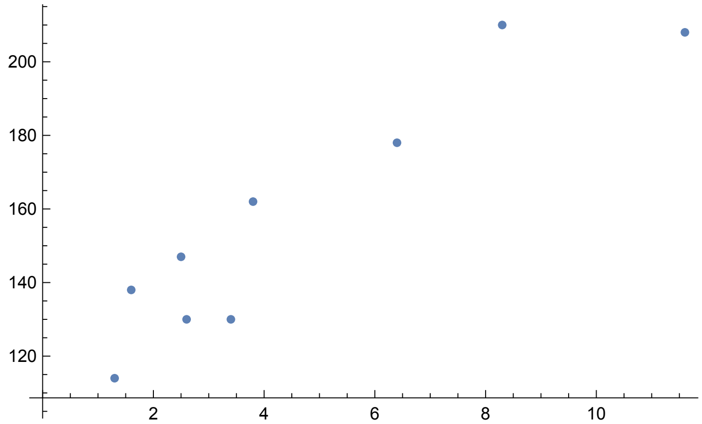
Umatilla	Morrow	Gilliam	Sherman	Wasco	Hood River	Portland	Columbia	Clatsop
2.5	2.6	3.4	1.3	1.6	3.8	11.6	6.4	8.3
147	130	130	114	138	162	208	178	210

Graphing the Data Set

```
In[27]:= data = Transpose[{index, deaths}];
```

In[28]:=
```
ListPlot[data]
```

Out[28]=



Finding the Least Squares Line

y = mx+b

$A = \sum_{i=1}^{n} (y_n)^2$

$B = \sum_{i=1}^{n} (x_n)^2$

$C = \sum_{i=1}^{n} x_n$

$D = \sum_{i=1}^{n} x_n * y_n$

$E = \sum_{i=1}^{n} y_n$

$m = \frac{Dn-EC}{Bn-C^2}$

$b = \frac{EB-DC}{Bn-C^2}$

In[29]:= $varA = \sum_{i=1}^{9} (deaths[[i]])^2;$

In[30]:= $varB = \sum_{i=1}^{9} (index[[i]])^2;$

In[31]:= $varC = \sum_{i=1}^{9} index[[i]];$

In[32]:= $varD = \sum_{i=1}^{9} (index[[i]] * deaths[[i]]);$

In[33]:= $varE = \sum_{i=1}^{9} (deaths[[i]]);$

In[34]:= $varm = \frac{(varD * 9 - varE * varC)}{varB * 9 - varC^2}$

Out[34]= 9.27386

In[35]:=
$$varb = \frac{(varE * varB - varC * varD)}{varB * 9 - varC^2}$$

Out[35]= 114.682

```
In[36]:= fitLine = varm * x + varb
```

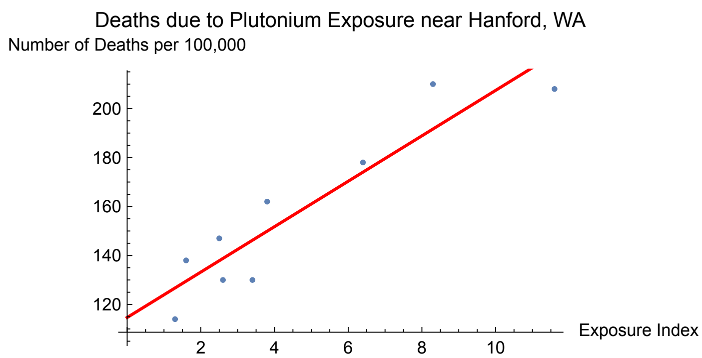
Out[36]= 114.682 + 9.27386 x

```
In[37]:= fitPlot = Plot[fitLine, {x, 0, 15}, PlotStyle → {Red}];
```

Graphing the Line of Best Fit

```
In[38]:= Show[ListPlot[data], fitPlot,
  PlotLabel → "Deaths due to Plutonium Exposure near Hanford, WA",
  AxesLabel → {"Exposure Index", "Number of Deaths per 100,000"}]
```

Out[38]=



Calculating Residuals

```
In[39]:= residuals = deaths - (varm * index + varb)
```

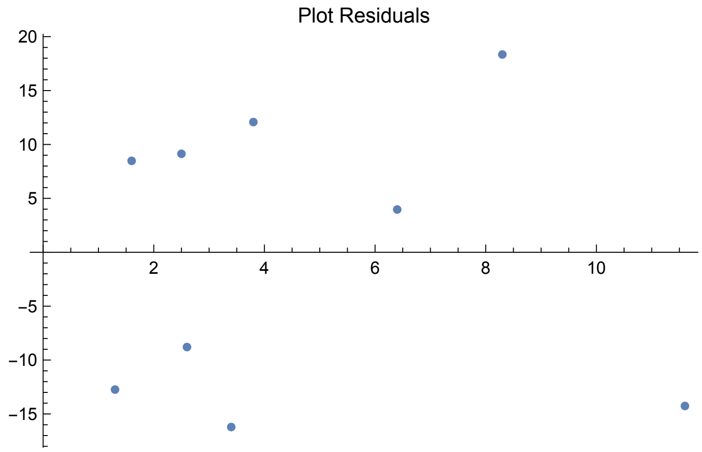
Out[39]= {9.13371, -8.79367, -16.2128, -12.7376, 8.48019, 12.0777, -14.2585, 3.96564, 18.3453}

```
In[40]:= resData = Transpose[{index, residuals}]
```

Out[40]= {{2.5, 9.13371}, {2.6, -8.79367}, {3.4, -16.2128}, {1.3, -12.7376}, {1.6, 8.48019}, {3.8, 12.0777}, {11.6, -14.2585}, {6.4, 3.96564}, {8.3, 18.3453}}

In[41]:=
```
ListPlot[resData, PlotLabel → "Plot Residuals"]
```

Out[41]=



Sum of Residuals

In[42]:=
$$resSum = \sum_{i=1}^{9} residuals[[i]]$$

Out[42]= 1.84741×10^{-13}

In[43]:=
```
reSSum = Total[residuals]
```

Out[43]= 1.84741×10^{-13}