**Question 1 – Background to Century National Bank**

The bank would like to know the characteristics of checking account customers. What is the balance of a typical customer? How many other bank services do the checking account customers use? Do the customers use the ATM service and, if so, how often are they used?

You are the head of the team and responsible for preparing the report. You select a random sample of 60 customers. In addition to the balance in each account at the end of last month, you determine: (1) the number of ATM (automatic teller machine) transactions in the last month; (2) the number of other bank services (a savings account, a certificate of deposit, etc.) the customer uses; (3) whether the customer has a debit card (this is a relatively new bank service in which charges are made directly to the customer’s account); and (4) whether or not interest is paid on the checking account. The sample includes customers from the branches in Cincinnati, Ohio; Atlanta, Georgia; Louisville, Kentucky; and Erie, Pennsylvania.

These data are contained in the data file CNB60.MTB and have the following variable definitions:

|  |  |
| --- | --- |
| Variable | Description |
| Balance | Account balance in $ |
| ATM | Number of ATM transactions for the month |
| Services | Number of other bank services used |
| Debit | Has a debit card (0 = no, 1 = yes) |
| Interest | Receives interest on the account (0 = no, 1 = yes) |
| City | City where banking is done (1=Cincinnati, 2=Atlanta, 3=Louisville, 4=Erie, PA) |

Refer to the description of Century Nation Bank in the Background section above. Using checking account balance as the response (Y) variable and using either the individual has a debit card variable OR whether interest is paid on the particular account as your predictor variable, write a report indicating how account balance relates to your predictor variable. Those seeking more adventure can do a 2-sample t-test since the debit card and interest variables are binary [each takes on two values]. How to the regression results compare to the 2-sample t-results. Check the p-values on your results using a significance level of α = 0.05.

Don’t forget question 2 that follows.**Question 2. Fun with regression. Brand New Question**

For the following regression data sets (4 of them), do the following activities in order. It is very important that you do each step in sequence. You can easily highlight the data table below and copy and paste to MINITAB.

1. Run the simple linear regressions and report the four estimated regression equations. The response variables are YA, YB, YC, and YD. The predictor variables are XA, XB, XC, and XD. Keep the pairs together (YA with XA and so on). You should be able to summarize the four regression equations that you obtained in a few sentences.
2. Do a scatterplot of each of the data sets. Do the scatter plots match your expectations based on part a above? Just be honest. A few sentences should be sufficient.
3. Do a plot of residuals for each of the data sets. Make comparisons between the scatterplot and residual plot for each model. Again a few sentences should suffice for each model.

The data set contains four pairs of X and Y values. Model 1 has variables XA and YA, Model 2 has variables XB and YB, and so on.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| YA | XA |  | YB | XB |  | YC | XC |  | YD | XD |
| 8.04 | 10 |  | 9.14 | 10 |  | 7.46 | 10 |  | 6.58 | 8 |
| 9.96 | 14 |  | 8.1 | 14 |  | 8.84 | 14 |  | 5.76 | 8 |
| 5.68 | 5 |  | 4.74 | 5 |  | 5.73 | 5 |  | 7.71 | 8 |
| 6.95 | 8 |  | 8.14 | 8 |  | 6.77 | 8 |  | 8.84 | 8 |
| 8.81 | 9 |  | 8.77 | 9 |  | 7.11 | 9 |  | 8.47 | 8 |
| 10.84 | 12 |  | 9.13 | 12 |  | 8.15 | 12 |  | 7.04 | 8 |
| 4.26 | 4 |  | 3.1 | 4 |  | 5.39 | 4 |  | 5.25 | 8 |
| 4.82 | 7 |  | 7.26 | 7 |  | 6.42 | 7 |  | 12.5 | 19 |
| 8.33 | 11 |  | 9.26 | 11 |  | 7.81 | 11 |  | 5.56 | 8 |
| 7.58 | 13 |  | 8.74 | 13 |  | 12.74 | 13 |  | 7.91 | 8 |
| 7.24 | 6 |  | 6.13 | 6 |  | 6.08 | 6 |  | 6.89 | 8 |