1.

From here consider .

Thus, . Thus, by the factorization theorem,

is a sufficient statistic.

2.

The joint density is then:

Essentially need to find a sort of T that fulfills This can be rearranged as: Thus, let

Then:

The factorization theorem holds for T. Thus, is a one-dimensional sufficient statistic for .

3.

As an example, showed in the Remarks on Big Data and Sufficient Statistics, it was already proven that for this problem is sufficient and thus is also sufficient. It follows that reordering the Xs is simply an application of the communicative property and it does not alter the results of and in any way. Then if we consider the definition of sufficiency, or the reordering would logically not depend on or meaning that the reordering is a sufficient statistic.

Regarding 4., I decided to put the output in the appendix which is basically my 4. For ease of reference and grading, I have put the individual answers with reasoning here as follows:

5. According to the SAS output and regression results, the monthly sales for pizza at this outlet is estimated to be **59220 pizzas**.

6. The regression output yields 12 degrees of freedom on the error term using the T-statistic. This means that the the 95% confidence interval for the error term consists of 2\*2.179 standard deviations.

Thus, the sales will differ by plus or minus **4879.761 pizzas**.

7. The normality tests in the SAS output reveal that the data is normally distributed. Thus, the 95% CI interval is the one listed in the SAS output: **(48587,69853)**.

8. Since all outlets will be producing the same, the mean estimate can come from the regression output. As such, the mean quantity sold is calculated via:

**59220.18 pizzas sold on average**

9.Again using similar logic to 6, the precise margin of error is found via:

Thus, I expect that my mean estimate will miss the actual mean estimate by plus or minus **2326.526 pizzas**.

10. While the central limit theorem does not directly apply here since the sample size is large, normality tests were run on the quantity sold earlier and the Shapiro–Wilk test revealed that the population amount of pizzas sold are normally distributed. Thus, the 95% confidence interval is: **(54151,64290)**.

11. Her opinion is within **2 sds** of the regression estimate based on the data. If we consider 2 sds to be consistent, then her opinion **is consistent with the data**.

12. With just considering the monthly advertising expenditure, the regression coefficient estimate is 1.75476 and it has a standard deviation of 0.18617 which **is consistent with the data** as her estimate falls within 2 sds of the estimate.

13. Under a 0.05 significance level, none of the normality tests’ null hypothesis that the residuals are normally distributed are rejected. **Thus, the residuals are normally distributed**.

| **Moments** | | | |
| --- | --- | --- | --- |
| **N** | 15 | **Sum Weights** | 15 |
| **Mean** | 63740 | **Sum Observations** | 956100 |
| **Std Deviation** | 13168.4037 | **Variance** | 173406857 |
| **Skewness** | -0.2936687 | **Kurtosis** | -0.5951811 |
| **Uncorrected SS** | 6.33695E10 | **Corrected SS** | 2427696000 |
| **Coeff Variation** | 20.6595603 | **Std Error Mean** | 3400.06723 |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 63740.00 | **Std Deviation** | 13168 |
| **Median** | 63200.00 | **Variance** | 173406857 |
| **Mode** | . | **Range** | 44800 |
|  |  | **Interquartile Range** | 15900 |

| **Tests for Location: Mu0=0** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 18.74669 | **Pr > |t|** | <.0001 |
| **Sign** | **M** | 7.5 | **Pr >= |M|** | <.0001 |
| **Signed Rank** | **S** | 60 | **Pr >= |S|** | <.0001 |

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.963692 | **Pr < W** | 0.7562 |
| **Kolmogorov-Smirnov** | **D** | 0.168104 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.038681 | **Pr > W-Sq** | >0.2500 |
| **Anderson-Darling** | **A-Sq** | 0.245248 | **Pr > A-Sq** | >0.2500 |

| **Quantiles (Definition 5)** | |
| --- | --- |
| **Level** | **Quantile** |
| **100% Max** | 85300 |
| **99%** | 85300 |
| **95%** | 85300 |
| **90%** | 79400 |
| **75% Q3** | 71400 |
| **50% Median** | 63200 |
| **25% Q1** | 55500 |
| **10%** | 42100 |
| **5%** | 40500 |
| **1%** | 40500 |
| **0% Min** | 40500 |

| **Extreme Observations** | | | |
| --- | --- | --- | --- |
| **Lowest** | | **Highest** | |
| **Value** | **Obs** | **Value** | **Obs** |
| 40500 | 2 | 71100 | 13 |
| 42100 | 15 | 71400 | 7 |
| 50800 | 4 | 77200 | 11 |
| 55500 | 14 | 79400 | 6 |
| 55600 | 9 | 85300 | 1 |

|  |  |
| --- | --- |
| **Number of Observations Read** | 16 |
| **Number of Observations Used** | 15 |
| **Number of Observations with Missing Values** | 1 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 2 | 2206867853 | 1103433926 | 59.96 | <.0001 |
| **Error** | 12 | 220828147 | 18402346 |  |  |
| **Corrected Total** | 14 | 2427696000 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Root MSE** | 4289.79552 | **R-Square** | 0.9090 |
| **Dependent Mean** | 63740 | **Adj R-Sq** | 0.8939 |
| **Coeff Var** | 6.73015 |  |  |

| **Parameter Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | Intercept | 1 | 817.61031 | 17704 | 0.05 | 0.9639 |
| **Average\_Price** | Average\_Price | 1 | -2617.59327 | 1189.74351 | -2.20 | 0.0481 |
| **Monthly\_Advertising\_Expenditures** | Monthly\_Advertising\_Expenditures | 1 | 1.69157 | 0.16608 | 10.19 | <.0001 |

| **Output Statistics** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Obs** | **Dependent Variable** | **Predicted Value** | **Std Error Mean Predict** | **95% CL Mean** | | **95% CL Predict** | | **Residual** |
| **1** | 85300 | 83889 | 2344 | 78783 | 88995 | 73238 | 94539 | 1411 |
| **2** | 40500 | 44737 | 2500 | 39291 | 50183 | 33920 | 55555 | -4237 |
| **3** | 61800 | 67669 | 1698 | 63969 | 71368 | 57616 | 77721 | -5869 |
| **4** | 50800 | 46232 | 2156 | 41536 | 50929 | 35772 | 56692 | 4568 |
| **5** | 60600 | 54410 | 1566 | 50997 | 57823 | 44460 | 64360 | 6190 |
| **6** | 79400 | 76855 | 2308 | 71826 | 81883 | 66241 | 87468 | 2545 |
| **7** | 71400 | 65395 | 1135 | 62923 | 67867 | 55727 | 75063 | 6005 |
| **8** | 70700 | 69364 | 1224 | 66697 | 72031 | 59644 | 79084 | 1336 |
| **9** | 55600 | 57616 | 1310 | 54762 | 60470 | 47843 | 67388 | -2016 |
| **10** | 70900 | 71901 | 1705 | 68186 | 75617 | 61843 | 81960 | -1001 |
| **11** | 77200 | 81799 | 2014 | 77410 | 86187 | 71473 | 92124 | -4599 |
| **12** | 63200 | 62175 | 1668 | 58540 | 65810 | 52146 | 72204 | 1025 |
| **13** | 71100 | 71481 | 1928 | 67279 | 75682 | 61233 | 81728 | -380.6818 |
| **14** | 55500 | 54588 | 2619 | 48882 | 60295 | 43637 | 65540 | 911.5834 |
| **15** | 42100 | 47990 | 1824 | 44015 | 51964 | 37833 | 58146 | -5890 |
| **16** | . | 59220 | 2327 | 54151 | 64290 | 48587 | 69853 | . |

|  |  |
| --- | --- |
| **Sum of Residuals** | 0 |
| **Sum of Squared Residuals** | 220828147 |
| **Predicted Residual SS (PRESS)** | 335277882 |





| **Moments** | | | |
| --- | --- | --- | --- |
| **N** | 15 | **Sum Weights** | 15 |
| **Mean** | 0 | **Sum Observations** | 0 |
| **Std Deviation** | 3971.57891 | **Variance** | 15773439.1 |
| **Skewness** | -0.0216223 | **Kurtosis** | -0.9284463 |
| **Uncorrected SS** | 220828147 | **Corrected SS** | 220828147 |
| **Coeff Variation** | . | **Std Error Mean** | 1025.45727 |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 0.0000 | **Std Deviation** | 3972 |
| **Median** | 911.5834 | **Variance** | 15773439 |
| **Mode** | . | **Range** | 12080 |
|  |  | **Interquartile Range** | 6783 |

| **Tests for Location: Mu0=0** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 0 | **Pr > |t|** | 1.0000 |
| **Sign** | **M** | 0.5 | **Pr >= |M|** | 1.0000 |
| **Signed Rank** | **S** | 4 | **Pr >= |S|** | 0.8469 |

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.943968 | **Pr < W** | 0.4349 |
| **Kolmogorov-Smirnov** | **D** | 0.124104 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.037361 | **Pr > W-Sq** | >0.2500 |
| **Anderson-Darling** | **A-Sq** | 0.27957 | **Pr > A-Sq** | >0.2500 |

| **Quantiles (Definition 5)** | |
| --- | --- |
| **Level** | **Quantile** |
| **100% Max** | 6190.086 |
| **99%** | 6190.086 |
| **95%** | 6190.086 |
| **90%** | 6005.279 |
| **75% Q3** | 2545.358 |
| **50% Median** | 911.583 |
| **25% Q1** | -4237.329 |
| **10%** | -5868.602 |
| **5%** | -5889.847 |
| **1%** | -5889.847 |
| **0% Min** | -5889.847 |

| **Extreme Observations** | | | |
| --- | --- | --- | --- |
| **Lowest** | | **Highest** | |
| **Value** | **Obs** | **Value** | **Obs** |
| -5889.85 | 15 | 1411.14 | 1 |
| -5868.60 | 3 | 2545.36 | 6 |
| -4598.73 | 11 | 4567.89 | 4 |
| -4237.33 | 2 | 6005.28 | 7 |
| -2015.74 | 9 | 6190.09 | 5 |

| **Missing Values** | | | |
| --- | --- | --- | --- |
| **Missing Value** | **Count** | **Percent Of** | |
| **All Obs** | **Missing Obs** |
| . | 1 | 6.25 | 100.00 |

|  |  |
| --- | --- |
| **Number of Observations Read** | 16 |
| **Number of Observations Used** | 15 |
| **Number of Observations with Missing Values** | 1 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 1 | 2117789777 | 2117789777 | 88.84 | <.0001 |
| **Error** | 13 | 309906223 | 23838940 |  |  |
| **Corrected Total** | 14 | 2427696000 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Root MSE** | 4882.51372 | **R-Square** | 0.8723 |
| **Dependent Mean** | 63740 | **Adj R-Sq** | 0.8625 |
| **Coeff Var** | 7.66005 |  |  |

| **Parameter Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | Intercept | 1 | -32655 | 10305 | -3.17 | 0.0074 |
| **Monthly\_Advertising\_Expenditures** | Monthly\_Advertising\_Expenditures | 1 | 1.75476 | 0.18617 | 9.43 | <.0001 |

| **Output Statistics** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Obs** | **Dependent Variable** | **Predicted Value** | **Std Error Mean Predict** | **95% CL Mean** | | **95% CL Predict** | | **Residual** |
| **1** | 85300 | 81054 | 2228 | 76241 | 85867 | 69459 | 92648 | 4246 |
| **2** | 40500 | 42449 | 2587 | 36860 | 48037 | 30512 | 54386 | -1949 |
| **3** | 61800 | 70174 | 1434 | 67077 | 73271 | 59181 | 81167 | -8374 |
| **4** | 50800 | 48941 | 2014 | 44591 | 53292 | 37532 | 60351 | 1859 |
| **5** | 60600 | 56311 | 1487 | 53100 | 59523 | 45285 | 67338 | 4289 |
| **6** | 79400 | 72806 | 1586 | 69381 | 76232 | 61716 | 83897 | 6594 |
| **7** | 71400 | 64910 | 1267 | 62173 | 67647 | 54013 | 75807 | 6490 |
| **8** | 70700 | 68946 | 1376 | 65972 | 71919 | 57987 | 79905 | 1754 |
| **9** | 55600 | 58768 | 1367 | 55816 | 61720 | 47815 | 69722 | -3168 |
| **10** | 70900 | 73859 | 1656 | 70282 | 77436 | 62721 | 84997 | -2959 |
| **11** | 77200 | 80352 | 2167 | 75670 | 85033 | 68812 | 91892 | -3152 |
| **12** | 63200 | 64910 | 1267 | 62173 | 67647 | 54013 | 75807 | -1710 |
| **13** | 71100 | 74210 | 1680 | 70580 | 77840 | 63055 | 85365 | -3110 |
| **14** | 55500 | 50170 | 1914 | 46036 | 54304 | 38841 | 61499 | 5330 |
| **15** | 42100 | 48240 | 2072 | 43763 | 52716 | 36781 | 59698 | -6140 |
| **16** | . | 55083 | 1560 | 51714 | 58453 | 44010 | 66156 | . |

|  |  |
| --- | --- |
| **Sum of Residuals** | 0 |
| **Sum of Squared Residuals** | 309906223 |
| **Predicted Residual SS (PRESS)** | 403774068 |







| **Moments** | | | |
| --- | --- | --- | --- |
| **N** | 15 | **Sum Weights** | 15 |
| **Mean** | 0 | **Sum Observations** | 0 |
| **Std Deviation** | 4704.90794 | **Variance** | 22136158.8 |
| **Skewness** | -0.0565283 | **Kurtosis** | -1.1032911 |
| **Uncorrected SS** | 309906223 | **Corrected SS** | 309906223 |
| **Coeff Variation** | . | **Std Error Mean** | 1214.80201 |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 0.00 | **Std Deviation** | 4705 |
| **Median** | -1709.84 | **Variance** | 22136159 |
| **Mode** | . | **Range** | 14968 |
|  |  | **Interquartile Range** | 7440 |

| **Tests for Location: Mu0=0** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 0 | **Pr > |t|** | 1.0000 |
| **Sign** | **M** | -0.5 | **Pr >= |M|** | 1.0000 |
| **Signed Rank** | **S** | 2 | **Pr >= |S|** | 0.9341 |

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.930739 | **Pr < W** | 0.2799 |
| **Kolmogorov-Smirnov** | **D** | 0.175187 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.081238 | **Pr > W-Sq** | 0.1915 |
| **Anderson-Darling** | **A-Sq** | 0.459272 | **Pr > A-Sq** | 0.2315 |

| **Quantiles (Definition 5)** | |
| --- | --- |
| **Level** | **Quantile** |
| **100% Max** | 6593.72 |
| **99%** | 6593.72 |
| **95%** | 6593.72 |
| **90%** | 6490.16 |
| **75% Q3** | 4288.50 |
| **50% Median** | -1709.84 |
| **25% Q1** | -3151.77 |
| **10%** | -6139.58 |
| **5%** | -8374.14 |
| **1%** | -8374.14 |
| **0% Min** | -8374.14 |

| **Extreme Observations** | | | |
| --- | --- | --- | --- |
| **Lowest** | | **Highest** | |
| **Value** | **Obs** | **Value** | **Obs** |
| -8374.14 | 3 | 4246.33 | 1 |
| -6139.58 | 15 | 4288.50 | 5 |
| -3168.17 | 9 | 5330.18 | 14 |
| -3151.77 | 11 | 6490.16 | 7 |
| -3110.09 | 13 | 6593.72 | 6 |

| **Missing Values** | | | |
| --- | --- | --- | --- |
| **Missing Value** | **Count** | **Percent Of** | |
| **All Obs** | **Missing Obs** |
| . | 1 | 6.25 | 100.00 |