**BUAN 6337 Homework 1\_Group 9**

**Question 1**

A gourmet pizza restaurant is considering adding new toppings to its menu. Each month they survey 10 customers about their preferences for three different toppings. The restaurant wants data on several different toppings. So they don’t always ask about the same three toppings. Customers rate each topping on a scale of 1(would never order) to 5 (would order often). The restaurant wants to compute average ratings for all toppings, so the ratings variables need to be numeric. The raw data file Pizza.csv has variables for the respondent’s ID, and the ratings for five different toppings: arugula, pine nuts, roasted butternut squash, shrimp, and grilled eggplant. The first two digits in the ID correspond to the month of the survey.

1. Examine the raw data file Pizza.csv and read it into SAS using the IMPORT procedure. Print the data set (on the results screen). Print a report that describes the contents of the data set to make sure all the variables are the correct type.
2. Open the raw data file in a simple editor like WordPad and compare the data values to the output from part b) to make sure that they were read correctly into SAS. In a comment in your report, identify any problems with the SAS data set that cannot be resolved using the IMPORT procedure. Explain what is causing the problem.
3. Read the same raw data file, Pizza.csv, this time using a DATA step (instead of the IMPORT procedure). Be sure to resolve any issues identified above.
4. Create a new dataset with the average ratings for each topping.

**Answer**

**a)**

The data file Pizza.csv has been uploaded into to SAS using the IMPORT procedure and below are images of the PROC print and PROC contents procedures to review the file contents. The variables are the correct type.

**Pizza table PROC print:**

Table

Description automatically generated

**Pizza PROC contents report:**

Table

Description automatically generated

**b)**

With PROC Import, SAS automatically scans the first 20 observations to detect the datatype. Missing data is represented by ‘.’ for numeric and blanks for character variables. Since first 20 rows are blanks for variables ‘Shrimp’ and ‘Eggplant’, they are identified instead as character variables which is resolved in the below section.

**Notepad view of Pizza raw data:**

Text

Description automatically generated

**c)**

The DATA step is easily created after using the IMPORT procedure by hitting “F4”. See the image below for the automatically created DATA step.

Text

Description automatically generated

**Table

Description automatically generated**

The data step and contents report both showed $ indicating characters next to the shrimp and eggplant variables. By editing the data step, we can now pull this data into SAS as numbers. See the edited data step below.

Text

Description automatically generated

The “.” in each cell instead of blanks for missing values, indicates that SAS now reads these as numeric values instead of character values.

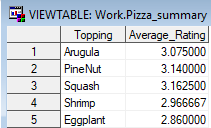
Graphical user interface, table

Description automatically generated

**d)**

The variables were renamed to show the correct names of “Topping” and “Average\_Rating”.

**Pizza Table:**



**Question 2**

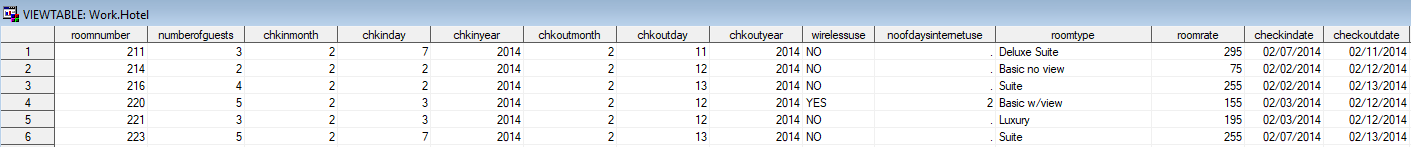
The new management of a local hotel decided to update their recently acquired (and very outdated) property by installing wireless Internet service for their guests. They are also considering updating their billing system because the method used by the previous owner seems faulty. In order to conduct a billing analysis, they would like some calculations about the guests who stayed with them during the first part of February (this was the first month after the change of ownership). The raw data file Hotel.dat contains variables with information on room number, number of guests, check-in month, day, year, check-out month, day, year, use of wireless Internet service, number of days of Internet use, room type, and room rate.

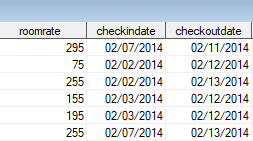
1. Examine the raw data file Hotel.dat and read it into SAS. Next, create date variables for the check-in and check-out dates, and format them to display as readable dates.
2. Create a variable that calculates the subtotal as the room rate times the number of days in the stay, plus a per person rate ($10 per day for each person beyond one guest, for example for 3 guests, the total per person rate will be (3-1)\*10=$20), plus an Internet service fee ($9.95 for a one-time activation and $4.95 per day of use).
3. Create a variable that calculates the grand total as the subtotal plus sales tax at 7.75%. The result should be rounded to two decimal places.
4. View the resulting data set. In a comment in your report, state the value for the grand total for room 211.

**Answer:**

**a)**

**Hotel table with check in and check out dates added and formatted as readable dates:**

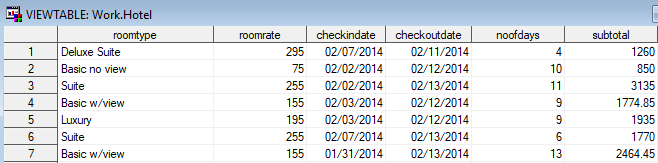




**b)**

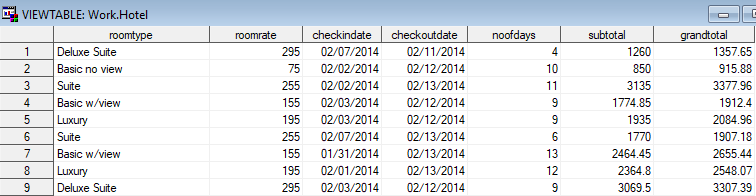
We made the assumption that the one-time activation fee is only charged when a guest uses the wireless internet service. A number of days variable “noofdays” was included to complete the daily rate calculation.

**Hotel table with number of days and subtotal added:**



**c)**

**Hotel table with grandtotal added:**



**d)**

The grand total for room 211 is **1357.65**.

**PROC Print of final Hotel table:**

