STAT 541: Final Exam

1. The file below contains information on graduate student applications. Save it as data set WORK.Applicants

Applicant	GRE-Q	GRE-V	GRE-W	TOEFL	LOR1	LOR2	LOR3
1CW	157	145	3.0	98	Davis	Chakraborty	
25Z	167	151	4.0	102	Xin	Wang	Zhu
QR5	166	155	4.5		Spiegelman		
AC8	159	150	3.5		Hoefner	Schultz	Riegel
T4Z	162	144	3.0	95	Singh	Shaked	Khaled

   (a) Use missing value formats so that when printing the data set, TOEFL scores are printed when available, the missing value for AC8 prints the message 'Not Received', and the missing value for QR5 prints the message 'N/A-US Citizen'.

   (b) Suppose T4Z retakes the TOEFL exam and earns a 98, and two more reference letters (Stigler and Arnold) arrive for QR5. Create a transactional data set containing these records.

   (c) Update the data set with the transactional data set. Undergraduates should use an UPDATE statement, while graduate students should use an index. Comment.

2. Referring to Question 1(a), show how you would update the TOEFL format in WORK.Applicants using MODIFY in PROC DATASETS.

3. The commands below create SAS data set AQChart with variables Ozone and PM_25_1 through PM_25_3. Ozone is ordered from 1 to 5 corresponding to observed levels of 0-50, 51-100, 101-150, 151-200, > 200. The three PM 2.5 variables correspond to levels of 0-50, 51-100 and > 100. The values in the chart represent air quality (1=Good, 2=Fair, 3=Poor).

```
data AQChart;
input ozone PM_25_1 PM_25_2 PM_25_3;
datalines;
1 1 1 2
2 1 2 2
3 2 2 3
4 2 3 3
5 3 3 3
;
run;
```

   (a) Read the tab-delimited text file AirQuality.txt into SAS; it includes air monitoring data from London. If the Date variable is giving you trouble, remove it from the text file and proceed with the rest of this problem.

(b) Create appropriate informats for Ozone and PM_25 in the AirQuality data set to code Ozone from 1 to 5 and to code PM_25 from 1 to 3.

(c) Use either PROC TRANSPOSE or match merge to assign an Air Quality Rating variable to each observation in AirQuality.txt. The printed data set should include text labels Good, Fair, and Poor as appropriate.

4. Return to AirQuality.txt. Use PICTURE to create a format for Date so that dates are printed as, e.g., 12PM Sunday, April 22, 2017. Demonstrate that your method works.

5. Our examples of NOTSORTED did not demonstrate that the method is actually effective. Using an unprotected workbook of your Fall08 data, sort the entire data set in Excel by Class (don't worry for now that the sort will be by alphabetical order).

(a) Import this data set into SAS and generate a table of MAJOR by CLASS, using the NOTSORTED option. Were the tables generated properly?

(b) Repeat the above, but use FORMATs or INFORMATs so that the classes are listed in proper order (Freshman, Sophomore, Junior, Senior).