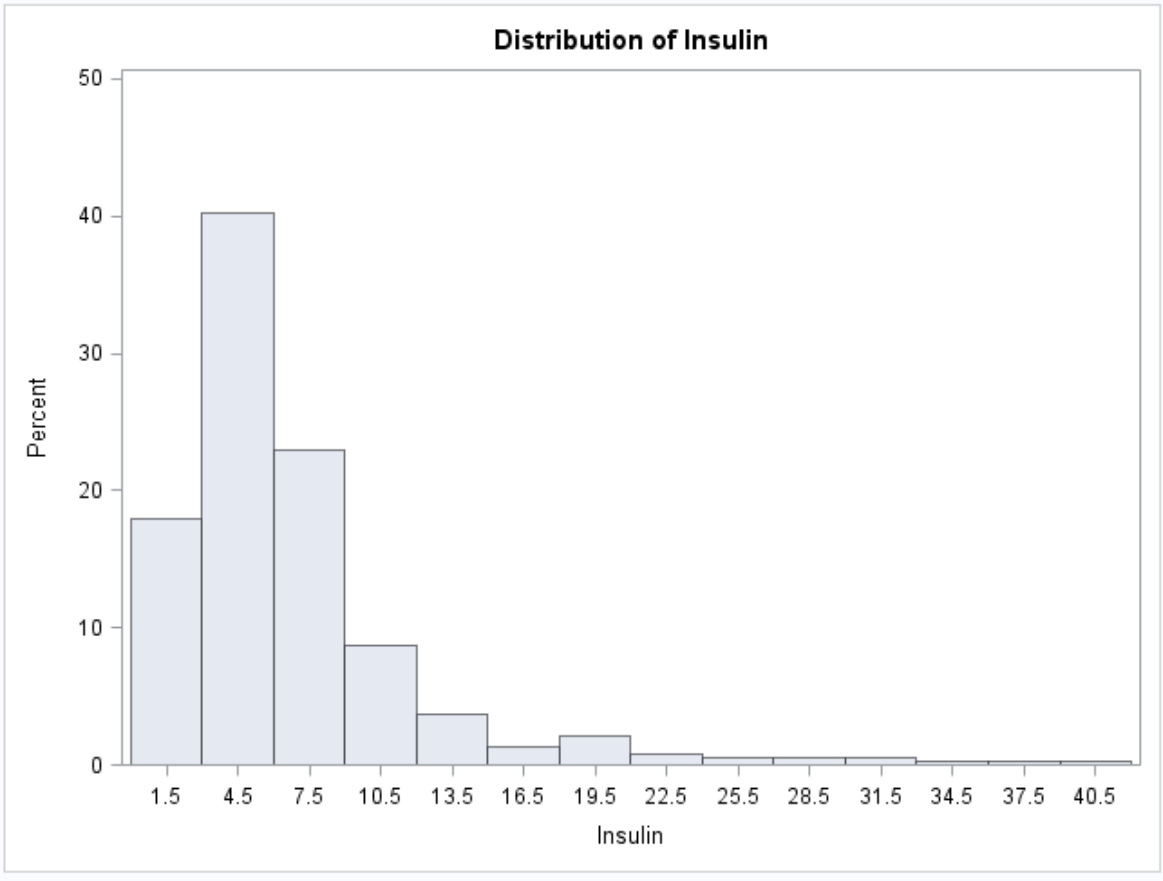
Name	Ian Towey
Student #	04128591
Course	STAT40840: Data Prog with SAS
Assignment #	2
Date	5th May 2016

1. **Dataset Summary**

Variable	N	N Miss	Mean	Std Error	t Value	Pr > |t|	Lower 95% CL for Mean	Upper 95% CL for Mean	Median	Std Dev	Variance	Minimum	Maximum	Range	Kurtosis	Skewness
ApoA1	397	17	142.9848866	1.7367267	82.33	<.0001	139.5705296	146.3992437	137.0000000	34.6040339	1197.44	13.0000000	275.0000000	262.0000000	2.0494873	0.4630786
ApoB	400	14	73.6500000	0.9099595	80.94	<.0001	71.8610858	75.4389142	72.0000000	18.1991903	331.2105263	32.0000000	128.0000000	96.0000000	0.0539492	0.2916124
ApoC2	400	14	3.5604500	0.0832238	42.78	<.0001	3.3968381	3.7240619	3.4000000	1.6644751	2.7704775	1.1000000	9.8900000	8.7900000	2.6099702	1.2600720
ApoC3	400	14	6.8467750	0.1276121	53.65	<.0001	6.5958989	7.0976511	6.8200000	2.5522418	6.5139382	0.6100000	15.4400000	14.8300000	1.6004459	0.2744690
ApoE	400	14	2.7880750	0.0672543	41.46	<.0001	2.6558580	2.9202920	2.5600000	1.3450856	1.8092552	0.6700000	6.8900000	6.2200000	0.6227797	0.8326187
Glucose	400	14	4.9463500	0.0260847	189.63	<.0001	4.8950694	4.9976306	4.9600000	0.5216940	0.2721646	3.2000000	6.9600000	3.7600000	1.1413666	0.3644317
NEFA	400	14	0.5349250	0.0140791	37.99	<.0001	0.5072465	0.5626035	0.5000000	0.2815825	0.0792887	0.0800000	1.7700000	1.6900000	2.0088805	1.1100445
Insulin	380	34	6.7548053	0.2837531	23.81	<.0001	6.1968776	7.3127329	5.2510000	5.5313671	30.5960215	0.7000000	40.1590000	39.4590000	10.2264133	2.7814552
Chol	400	14	4.4138250	0.0466615	94.59	<.0001	4.3220918	4.5055582	4.2200000	0.9332309	0.8709199	0.6900000	6.6200000	5.9300000	0.9128051	-0.1385235
TAG	400	14	1.2780250	0.0316593	40.37	<.0001	1.2157851	1.3402649	1.1000000	0.6331865	0.4009252	0.3400000	4.3600000	4.0200000	3.4817586	1.5959678

The table above presents summary statistics for the blood biochemcial component measured in the experiment. A few interesting points from this are

a) The standard deviation of the *Insulin* is very high
b) Insulin is right skewed with a few extreme value 6 standard deviations from the mean

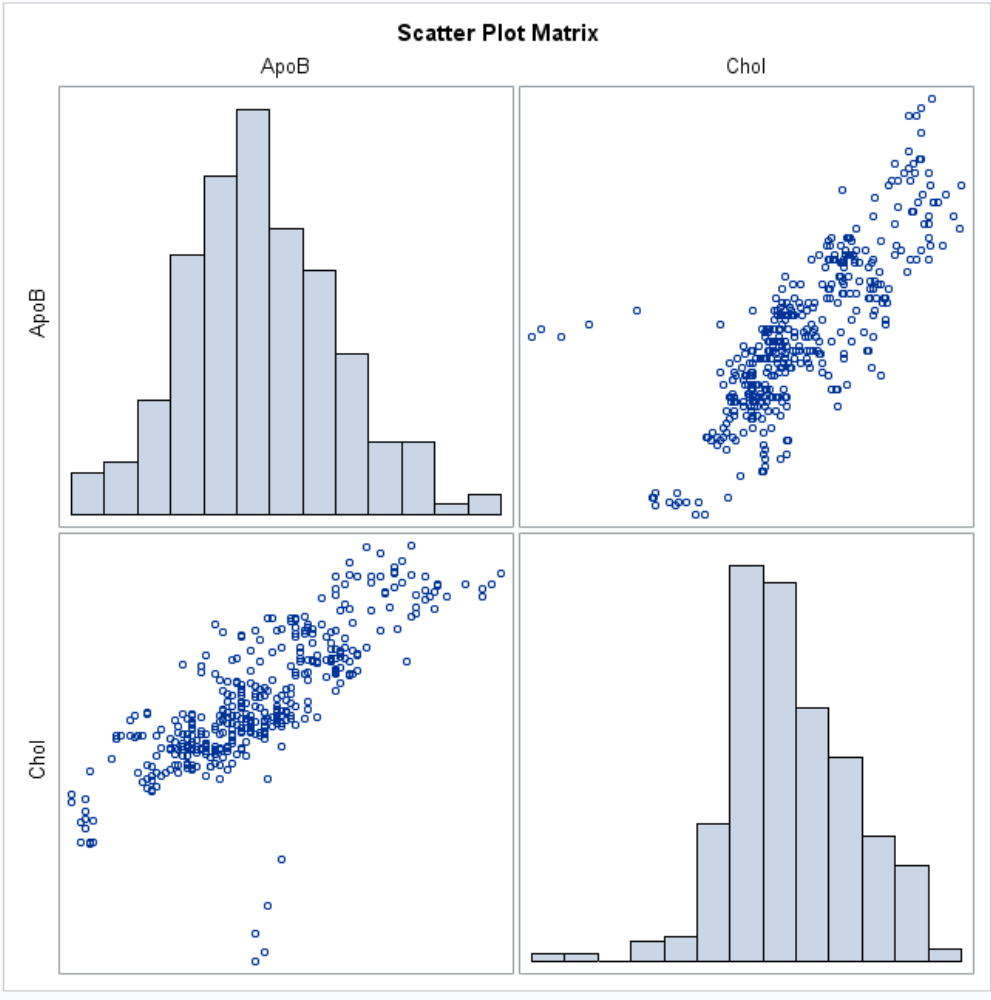


c)
d) Cholesterol is the only vaiable that is left skewed.

2. **Correlation Analysis**

Pearson Correlation Coefficients, N = 375 Prob > |r| under H0: Rho=0										
	ApoA1	ApoB	ApoC2	ApoC3	ApoE	Glucose	NEFA	Insulin	Chol	TAG
ApoA1	1.00000	0.00744 0.8859	-0.04817 0.3523	0.11863 0.0216	0.02485 0.6315	-0.03681 0.4773	0.03411 0.5101	0.02542 0.6236	0.25983 <.0001	0.01852 0.7208
ApoB	0.00744 0.8859	1.00000	0.39111 <.0001	0.40875 <.0001	0.12196 0.0181	0.10779 0.0369	0.01566 0.7625	0.01685 0.7450	0.78506 <.0001	0.24542 <.0001
ApoC2	-0.04817 0.3523	0.39111 <.0001	1.00000	0.65406 <.0001	0.37365 <.0001	0.15662 0.0024	-0.11647 0.0241	0.13199 0.0105	0.34900 <.0001	0.55376 <.0001
ApoC3	0.11863 0.0216	0.40875 <.0001	0.65406 <.0001	1.00000	0.29848 <.0001	0.16903 0.0010	-0.03985 0.4417	0.18553 0.0003	0.39843 <.0001	0.48969 <.0001
ApoE	0.02485 0.6315	0.12196 0.0181	0.37365 <.0001	0.29848 <.0001	1.00000	-0.01358 0.7932	-0.06798 0.1890	0.05463 0.2914	0.27977 <.0001	0.26690 <.0001
Glucose	-0.03681 0.4773	0.10779 0.0369	0.15662 0.0024	0.16903 0.0010	-0.01358 0.7932	1.00000	-0.08296 0.1087	0.35238 <.0001	0.12754 0.0134	0.16878 0.0010
NEFA	0.03411 0.5101	0.01566 0.7625	-0.11647 0.0241	-0.03985 0.4417	-0.06798 0.1890	-0.08296 0.1087	1.00000	-0.33851 <.0001	0.03661 0.4797	0.05039 0.3304
Insulin	0.02542 0.6236	0.01685 0.7450	0.13199 0.0105	0.18553 0.0003	0.05463 0.2914	0.35238 <.0001	-0.33851 <.0001	1.00000	0.03528 0.4958	0.10254 0.0472
Chol	0.25983 <.0001	0.78506 <.0001	0.34900 <.0001	0.39843 <.0001	0.27977 <.0001	0.12754 0.0134	0.03661 0.4797	0.03528 0.4958	1.00000	0.30143 <.0001
TAG	0.01852 0.7208	0.24542 <.0001	0.55376 <.0001	0.48969 <.0001	0.26690 <.0001	0.16878 0.0010	0.05039 0.3304	0.10254 0.0472	0.30143 <.0001	1.00000

a) Apolipoprotein B and Cholesterol has a high positive correlation at 0.78506



Apolipoprotein C2 and Apolipoprotein C3 has a high positive correlation at 0.65406

b) Apolipoprotein A1 and Apolipoprotein B have least correlation

3. **Partial Correlation Analysis**

1 Partial Variables:	BMI
2 Variables:	Glucose Insulin

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Partial Variance	Partial Std Dev
BMI	378	24.60291	2.58895	9300	19.60000	33.30000		
Glucose	378	4.95156	0.52317	1872	3.20000	6.96000	0.26384	0.51365
Insulin	378	6.77963	5.53541	2563	0.70000	40.15900	29.83103	5.46178

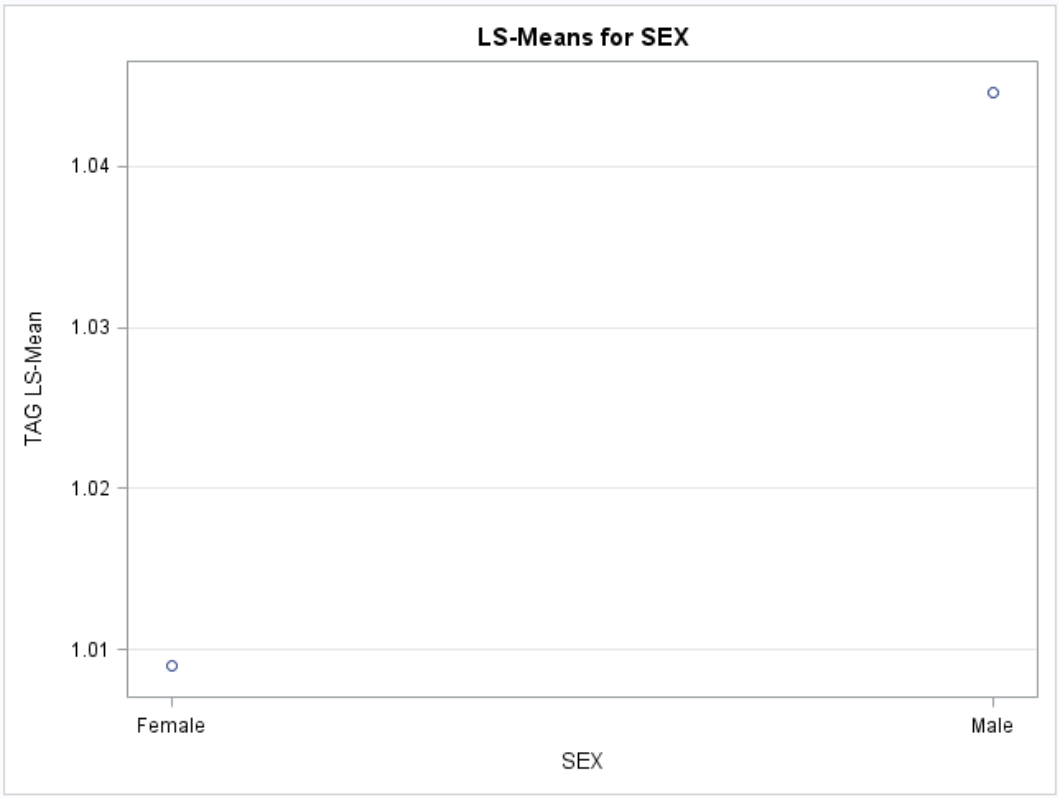
Pearson Partial Correlation Coefficients, N = 378
Prob > |r| under H0: Partial Rho=0

	Glucose	Insulin
Glucose	1.00000	0.32733 <.0001
Insulin	0.32733 <.0001	1.00000

a) The partial correlation between the variables Glucose and Insulin is 0.32733, which is a little less than the unpartialled correlation, 0.35238. The p-value for the partial correlation is less than 0.0001.

4. **Linear Model**

SEX Least Squares Means				
SEX	Estimate	Standard Error	z Value	Pr > |z|
Female	1.0090	0.07840	12.87	<.0001
Male	1.0446	0.08450	12.36	<.0001

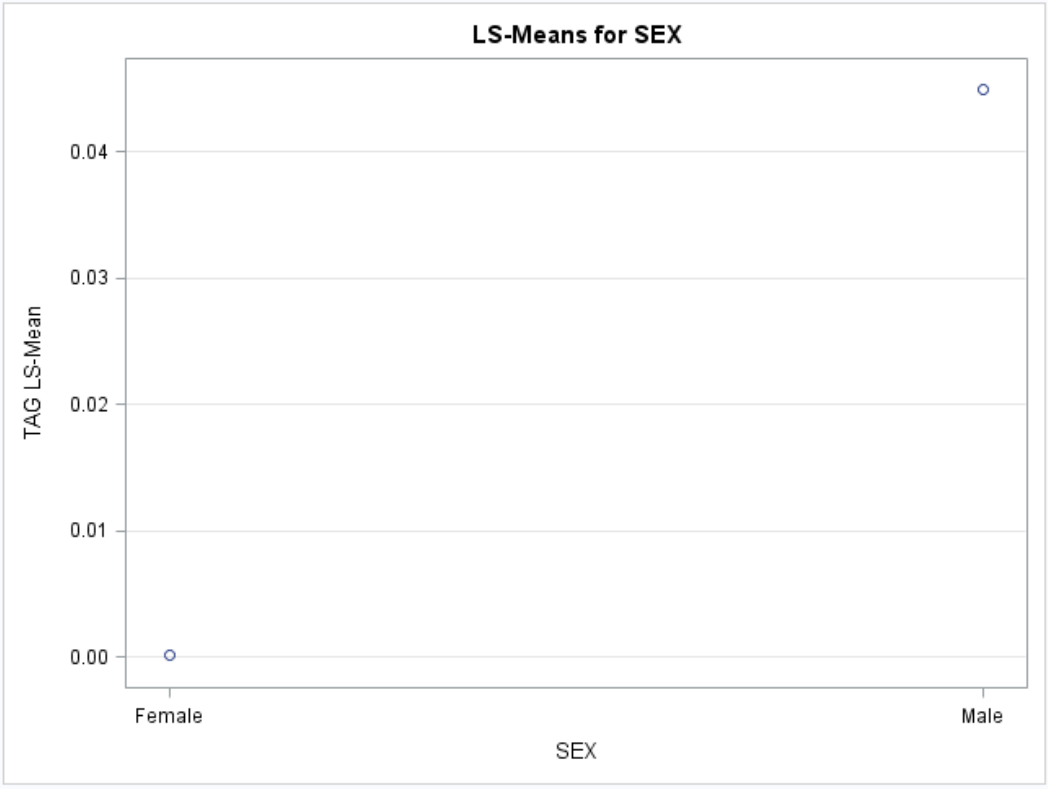


Differences of SEX Least Squares Means					
SEX	_SEX	Estimate	Standard Error	z Value	Pr > |z|
Female	Male	-0.03561	0.1171	-0.30	0.7609

a) The estimate for the TAG response variable are not different for males and females

5. **Log normal Model**

SEX Least Squares Means				
SEX	Estimate	Standard Error	z Value	Pr > |z|
Female	0.000190	0.07896	0.00	0.9981
Male	0.04492	0.08072	0.56	0.5779



Differences of SEX Least Squares Means					
SEX	_SEX	Estimate	Standard Error	z Value	Pr > |z|
Female	Male	-0.04473	0.1140	-0.39	0.6948

a) The estimate for the TAG response variable are not different for males and females for a log normal model

6. **Mixed linear model**

Solution for Fixed Effects						
Effect	SEX	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-0.5986	3.3309	60	-0.18	0.8580
AGE		-0.05978	0.02841	60	-2.10	0.0396
SEX	Female	0.3198	0.6461	60	0.49	0.6224
SEX	Male	0	.	.	.	.
BMI		0.3602	0.1249	60	2.88	0.0054

a) ar(1) model

Solution for Fixed Effects						
Effect	SEX	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-1.8035	1.9222	60	-0.94	0.3519
AGE		-0.04010	0.01639	60	-2.45	0.0174
SEX	Female	0.5580	0.3727	60	1.50	0.1396
SEX	Male	0	.	.	.	.
BMI		0.2954	0.07208	60	4.10	0.0001

b) arh(1) model

The fit of arh(1) model has a smaller variance than that of the ar(1) model