Data Customer\_X;

Input Gender $ Age Region $11.;

Datalines;

F 34 Portsmouth

F 45 Southampton

M 36 Leeds

M 65 Durham

M 56 Bristol

M 19 Newcastle

F 23 London

;

Data Customer\_X;

Input ID Class $ Height $ Weight $ Football Basketball Hockey;

Datalines;

1 A Over5.7 Above50 1 0 1

2 A Over5.7 Above50 1 1 0

3 B Over5.7 Below50 1 1 .

4 B Under5.7 Below50 1 1 1

5 A Over5.7 Below50 1 1 1

6 A Over5.7 Above50 1 . 1

;

Proc Print;

Run;

Data Class;

Input ID Class $ Height $ Weight $;

Datalines;

1 A Over5.7 Above50 1 0 1

2 A Over5.7 Above50 1 1 0

3 B Over5.7 Below50 1 1 .

4 B Under5.7 Below50 1 1 1

5 A Over5.7 Below50 1 1 1

6 A Over5.7 Above50 1 . 1

;

Proc Freq Data = Class;

Run;

Proc Freq Data = Class;

Table Class;

Run;

Proc Freq Data = Class;

Table Class\*Height;

Run;

Proc Freq Data=Class;

Table Class\*Height/nocol norow nocum nofreq;

Run;

Proc Freq Data=Class;

Table Class\*Weight\*Height/nocol norow nocum nofreq;

Run;

Proc Freq Data=Class;

Table Class Weight\*Height/nocol norow nocum nofreq;

Run;

Ods graphics on;

Proc Freq Data=Class;

Table Class Weight\*Height/plots=freqplot;

Run;

Data SportsInjury;

Input Activity $ Warmup Injury Count;

Datalines;

Running 1 0 5

Running 0 1 15

Running 1 1 3

Football 1 0 16

Football 1 1 10

Football 1 1 4

Squash 1 0 2

Squash 0 1 10

Squash 1 1 0

Weights 1 0 12

Weights 0 1 6

Weights 1 1 3

Others 0 0 10

;

Proc Sql;

Create table Analyse as

Select Warmup, Injury, Sum(Count) as Cases

From SportsInjury

Group by 1,2;

Quit;

Proc Print Noobs;

Run;

Proc Sort Data = Analyse;

By Warmup descending Injury;

Run;

Proc Freq Data=Analyse Order=Data;

Tables Warmup\*Injury / chisq relrisk;

Weight Cases;

Run;

ODS Select Frequencies;

Proc Univariate Data = Analyse Freq;

Var \_All\_;

Run;

Data Transactions;

Input CustId $ Credit;

Datalines;

A2112 234

A2342 532

A2345 345

A6345 234

B3234 234

B6345 456

C465A 675

D4436 790

E4603 645

F0945 709

F435F 999

H0032 009

;

ODS Select BasicMeasures Extremeobs;

Proc Univariate Data = Transactions;

Var Credit;

Run;

ods select ParameterEstimates GoodnessOfFit FitQuantiles Bins MyPlot;

ods graphics off;

Proc Univariate Data = Transactions;

Histogram Credit / normal

name='MyPlot';

inset n normal(ksdpval) / pos = ne format = 6.3;

run;

ODS Select TestsforLocation LocationCounts;

Proc Univariate Data = Transactions MU0=200 Loccount;

Var Credit;

Run;

Proc Means Data = Transactions;

Run;

Proc Means Data = SportsInjury;

Class Warmup Injury;

Var Count;

Run;

Proc Sort Data = Customer\_X;

By Class;

Run;

Proc Means Data = Customer\_X;

By Class;

Class Height;

Var Basketball;

Run;

Proc Summary Data = Customer\_X;

By Class;

Class Height;

Var Basketball;

Output Out=Test n=n mean=mean std=stdev min=min max=max;

Run;

Proc Print Data=\_LAST\_ (Drop = \_TYPE\_ Rename=(\_FREQ\_=Nobs)) Noobs;

Where Height ne "";

Run;

Data Model;

Set Import;

Run;

Proc Corr Data = Model;

ID Date;

With Stock;

Var Basket\_Index -- M1\_Money\_Supply\_Index;

Run;

Data Build Validation;

Set Model;

If Date lt '01Dec2017'd then output Build;

Else output Validation;

Run;

PROC REG DATA=build plots=diagnostics(unpack);

ID date;

MODEL stock = basket\_index -- m1\_money\_supply\_index;

RUN;

PROC REG DATA=build OUTEST=REGOUT;

ID DATE;

MODEL Stock = basket\_index eps p\_e\_ratio global\_mkt\_share media\_analytics\_index

m1\_money\_supply\_index top\_10\_gdp;

RUN;

PROC SCORE DATA=validation

SCORE=REGOUT OUT=RSCOREP TYPE=PARMS;

var basket\_index eps p\_e\_ratio global\_mkt\_share media\_analytics\_index

m1\_money\_supply\_index top\_10\_gdp;

RUN;

Proc Print Data = RSCOREP (Keep = Date Stock Model1 Rename = (Model1=Predicted\_Stock\_Value));

Run;

Proc Transpose Data = Class Out = Transposed\_Class;

Run;

Proc Print;

Run;

proc transpose data=Class out=idnumber name=Test

prefix=sn;

id class;

run;

Data Base;

Input CustID Year Avg\_Credit Avg\_Debit Spend\_Indicator $;

Datalines;

1010 16 235 245 R

1010 17 230 220 A

1010 18 235 200 G

1010 19 254 220 G

1011 16 653 650 A

1011 17 650 610 G

1011 18 640 620 G

1011 19 650 656 A

1012 16 569 569 R

1012 17 560 550 G

1012 18 550 550 R

1012 19 450 400 G

;

Proc Print;

Run;

Data Base\_Narrow (Keep = CustID Year);

Set Base;

Where CustID=1010;

Run;

Proc Print Noobs;

Run;

Proc Transpose Data = Base\_Narrow Out=Wide;

Run;

Proc Print Noobs;

Run;

Proc Transpose Data = Base\_Narrow Out=Wide Prefix=ID Prefix=Year Prefix=Avg\_Credit;

Run;

Proc Print Noobs;

Run;

Data Base\_Narrow (Keep = CustID Year Avg\_Credit);

Set Base;

Run;

Proc Print Noobs;

Run;

Proc Transpose Data = Base\_Narrow Out=Wide (Drop=\_NAME\_) Prefix=Year;

By CustID;

ID Year;

Var Avg\_Credit;

Run;

Title Height = 8pt "Average Credit of Customers Across Years";

Proc Print Noobs;

Run;