STAT 615 (Baron)

Data entry, Print, plot, simple stats

Refers to the datafile "pulse.dat" (in Unix). In Windows, write the full path.

Reduce the size of plots to 33% (option)

plots two graphs on the same picture

Options verifying Normality

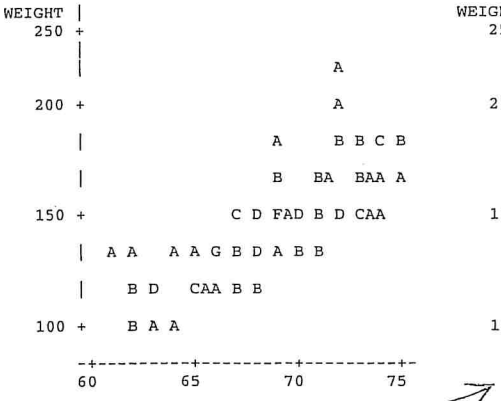
```
DATA;
INFILE pulse;
INPUT ROW PULSE1 PULSE2 RAN SMOKES SEX HEIGHT WEIGHT ACTIVITY;

PROC PRINT;

PROC PLOT HPERCENT=33 VPERCENT=33;
PLOT WEIGHT*HEIGHT;
PLOT WEIGHT*HEIGHT=SEX;
PLOT PULSE1*WEIGHT='*' PULSE2*WEIGHT='o' / OVERLAY;

PROC MEANS;
VAR PULSE1 PULSE2 SMOKES SEX HEIGHT WEIGHT;

PROC UNIVARIATE NORMAL PLOT;
VAR HEIGHT;

RUN;
```

PROC PRINT

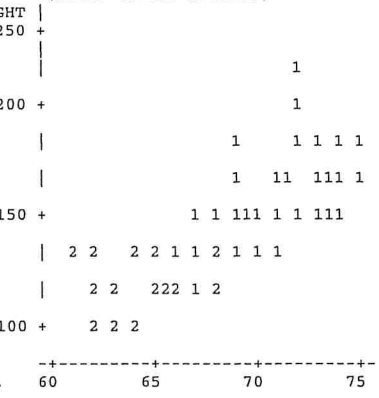
Obs	ROW	PULSE1	PULSE2	RAN	SMOKES	SEX	HEIGHT	WEIGHT	ACTIVITY
1	1	64	88	1	2	1	66.00	140	2
2	2	58	70	1	2	1	72.00	145	2
3	3	62	76	1	1	1	73.50	160	3
.....									
92	92	76	76	2	2	2	61.75	108	2

PROC PLOT

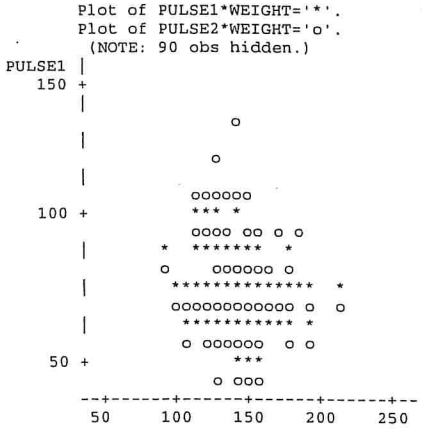
Plot of WEIGHT*HEIGHT. A=1, B=2, etc.



Plot of WEIGHT*HEIGHT=SEX. (NOTE: 48 obs hidden.)



Plot of PULSE1*WEIGHT='*'. Plot of PULSE2*WEIGHT='o'. (NOTE: 90 obs hidden.)



PROC PLOT

Plots "1" for men "2" for women

The MEANS Procedure

PROC MEANS

Variable	N	Mean	Std Dev	Minimum	Maximum
PULSE1	92	72.8695652	11.0087052	48.0000000	100.0000000
PULSE2	92	80.0000000	17.0937943	50.0000000	140.0000000
SMOKES	92	1.6956522	0.4626519	1.0000000	2.0000000
SEX	92	1.3804348	0.4881540	1.0000000	2.0000000
HEIGHT	92	68.7173913	3.6592907	61.0000000	75.0000000
WEIGHT	92	145.1521739	23.7393978	95.0000000	215.0000000

$\frac{\sum(X-\bar{X})^3/n}{S^3}$ measures symmetry

$\bar{X}$ S $\sum(X_i-\bar{X})^2$

The UNIVARIATE Procedure
Variable: HEIGHT

PROC UNIVARIATE

Moments

N	92	Sum Weights	92
Mean	68.7173913	Sum Observations	6322
Std Deviation	3.65929071	Variance	13.3904085
Skewness	-0.2246482	Kurtosis	-0.8044121
Uncorrected SS	435649.875	Corrected SS	1218.52717
Coeff Variation	5.32513042	Std Error Mean	0.38150743

$\sum X_i^2$

$\hat{\gamma} = \frac{S}{\mu} \cdot 100\%$

$S/\sqrt{n}$

Basic Statistical Measures

Location		Variability	
Mean	68.71739	Std Deviation	3.65929
Median	69.00000	Variance	13.39041
Mode	68.00000	Range	14.00000
		Interquartile Range	6.00000

(max - min)

(Q3 - Q1)

$\frac{\sum(X-\bar{X})^4/n}{S^4} - 3$ measures deviation from Normality