1.

```
PROC IMPORT DATAFILE="sashelp.heart" OUT=heart;
RUN;

PROC TTEST DATA=heart H0=150;
  VAR Weight;
RUN;

PROC UNIVARIATE DATA=heart;
  VAR Weight;
  NORMAL PLOT;
RUN;

PROC UNIVARIATE DATA=heart;
  VAR Weight;
  HISTOGRAM;
RUN;

PROC SGPLOT DATA=heart;
  BOXPLOT Weight / CATEGORY=1;
RUN;
```

About the tests of normality rejecting the null hypothesis at the .05 level, it indicates evidence against the normality assumption. However, it is important to consider the sample size and the robustness of the t-test to deviations from normality. If the sample size is large (typically above 30), the t-test can still be reliable even if the normality assumption is violated. Additionally, the shape of the histogram and box plot can provide visual insights into the distribution of the data. It would be best to interpret the results of the t-test in conjunction with the visual representations and the specific context of your analysis.

2.

```
PROC SORT DATA=sashelp.fish OUT=fish_sorted;
  BY Species;
RUN;

PROC TTEST DATA=fish_sorted(where=(Species='Smelt')) H0=10;
```

```
  VAR Weight;
RUN;

PROC UNIVARIATE DATA=fish_sorted(where=(Species='Smelt'));
  VAR Weight;
  EXACT WILCOXON / ONEPOPULATION(MU=10);
RUN;
```

By running the above code, SAS will provide the results of both the parametric one-sample t-test and the nonparametric one-sample Wilcoxon signed-rank test for the Smelt species in the "Fish" dataset. To compare the parametric and nonparametric tests, you can look at their respective p-values. If both tests yield p-values greater than the significance level (e.g., $\alpha = 0.05$), it suggests that there is not enough evidence to reject the null hypothesis. If the p-values are less than the significance level, it indicates evidence to reject the null hypothesis. However, it's important to note that the conclusions drawn from the two tests may differ. The parametric test assumes normality and the nonparametric test does not, so they can lead to different outcomes. If the assumptions of the parametric test are not met, the nonparametric test can provide a more reliable inference. To reach the same conclusion using these tests, both the parametric and nonparametric tests should yield p-values greater than the significance level, indicating that there is not enough evidence to reject the null hypothesis that the mean weight of Smelt is equal to 10.

3.

1.

```
DATA First10;
  SET SASHELP.AIR (OBS=10);
RUN;
PROC PRINT DATA=First10;
RUN;
```

2.

```
PROC MEANS DATA=SASHELP.AIR VAR Air;
  OUTPUT OUT=SummaryStats MEAN=Mean MIN=Min MAX=Max;
RUN;
PROC UNIVARIATE DATA=SASHELP.AIR;
  VAR Air;
```

```
  HISTOGRAM;
RUN;
PROC SGPLOT DATA=SASHELP.AIR;
  BOXPLOT Air / VBOX;
RUN;
```

3.

```
PROC TTEST DATA=SASHELP.AIR H0=285;
  VAR Air;
RUN;
PROC NPAR1WAY DATA=SASHELP.AIR WILCOXON;
  VAR Air;
  EXACT;
RUN;
```

4.

```
data difference;
  call streaminit(13579);
  do Subj = 1 to 20;
    Diff = .6 - rand('uniform');
    output;
  end;
run;

PROC TTEST DATA=difference H0=0;
  VAR Diff;
RUN;
```

5.

```
data difference;
  call streaminit(13579);
  do Subj = 1 to 200;
    Diff = .6 - rand('uniform');
    output;
  end;
run;
```

```
PROC TTEST DATA=difference H0=0;
  VAR Diff;
RUN;

PROC UNIVARIATE DATA=difference;
  VAR Diff;
  NORMAL PLOT;
RUN;
```

By including the PROC UNIVARIATE step with the NORMAL option, SAS will perform a Shapiro-Wilk test for normality and provide a normal probability plot.
Now, let's analyze the results. If the p-value from the Shapiro-Wilk test is greater than the significance level (e.g., $\alpha = 0.05$), it suggests that there is no significant evidence to reject the null hypothesis of normality. However, if the p-value is less than the significance level, it indicates evidence against the normality assumption.
Based on the output, if the p-value from the Shapiro-Wilk test is greater than 0.05 (or your chosen significance level), it suggests that the difference scores are approximately normally distributed. In this case, it would be reasonable to proceed with the t-test, even if the data does not exactly follow a normal distribution. The t-test is robust to deviations from normality, especially with large sample sizes.
Comparing the p-value obtained from the t-test between the initial dataset (with 20 observations) and the updated dataset (with 200 observations), the p-value is likely to decrease with the larger sample size. A larger sample size often leads to increased statistical power, making it more likely to detect even small deviations from the null hypothesis.