```
options nodate pageno=1; run;
ods rtf file ='C:\Users\Richard\Documents\Classes\Stat 5100\2017 Spring STAT
5100 001\Concord Summary.rtf';
ods graphics on;
data concord1;
    infile 'C:\Users\Richard\Documents\Classes\Stat 5100\2017 Spring STAT 5100
001\concord1.dat';
    input Household Water81 Water80 Water79 Income
          Education Retired People81 ChPeople People80;
    label Water81 = 'Household Water Consumption in 1981 (in cubic feet)';
    label Water80 = 'Household Water Consumption in 1980 (in cubic feet)';
    label Water79 = 'Household Water Consumption in 1979 (in cubic feet)';
    label People81 = 'Number of people in household in 1981';
    label People80 = 'Number of people in household in 1980';
    label ChPeople = 'Change in number of people in household 1980 to 1981';
    label Retired = 'Head of household is retired (1 = yes, 0 = no)';
    label Income1 = 'Household income in 1981';
    label Education = 'Years of Education of Head of Household';
run;

title1 "VARIABLES IN CONCORD WATER DATASET";
proc contents data=concord1;
run;

title1 "SUMMARY STATISTICS AND PLOTS FOR CONCORD WATER DATA FOR 1981";
proc univariate normal plots data=concord1;
    var Water81;
      title 'Histograms for 1981 Water Consumption';
      histogram Water81 / midpoints = 500 1500 2500 3500 4500 5500 6500 7500
                                      8500 9500 10500 normal(color=blue);
      histogram Water81 / midpoints = 250 750 1250 1750 2250 2750 3250 3750
                                      4250 4750 5250 5750 6250 6750 7250
                                      7750 8250 8750 9250 9750 10250
                                      normal(color=blue);
run;

data concord2;
    set concord1;
    logWater81 = log(Water81);
    label logWater81 = 'Log of Water Consumption in 1981';
    srtWater81 = sqrt(Water81);
    label srtWater81 = 'Square Root of Water Consumption in 1981';
    crtWater81 = Water81**(1/3);
    label crtWater81 = 'Cube Root of Water Consumption in 1981';
run;
```

```sas
title1 "SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION";
proc univariate normal plots data=concord2;
    var srtWater81 crtWater81 logWater81;
    histogram crtWater81 / normal;
run;

title1 "SCATTER PLOT OF 1981 WATER CONSUMPTION AGAINST 1980 WATER
CONSUMPTION";
proc sgplot data=concord1;
    scatter X=Water80 Y=Water81 / Group=Retired;
run;

proc sort data=concord2;
    by water81;
run;

title1 "Plot of Cube Root of 1981 Water Consumption against 1981 Water
Consumption";
proc sgplot data=concord2;
    series X=Water81 Y=crtWater81;
run;

title1 'Correlations among Variable in Concord Water Data';
proc corr data=concord1 plots(maxpoints=200000)=matrix(nvar=all histogram);
    var Water81 Water80 Water79 Income
        Education Retired People81 ChPeople People80;
run;

ods graphics off;
ods rtf close;

quit;
```

# VARIABLES IN CONCORD WATER DATASET

## The CONTENTS Procedure

| Data Set Name | WORK.CONCORD1 | Observations | 496 |
|---|---|---|---|
| Member Type | DATA | Variables | 10 |
| Engine | V9 | Indexes | 0 |
| Created | 01/11/2017 10:46:28 | Observation Length | 80 |
| Last Modified | 01/11/2017 10:46:28 | Deleted Observations | 0 |
| Protection | | Compressed | NO |
| Data Set Type | | Sorted | NO |
| Data Representation | WINDOWS_64 | | |
| Encoding | wlatin1   Western (Windows) | | |

| Engine/Host Dependent Information | |
|---|---|
| Data Set Page Size | 65536 |
| Number of Data Set Pages | 1 |
| First Data Page | 1 |
| Max Obs per Page | 817 |
| Obs in First Data Page | 496 |
| Number of Data Set Repairs | 0 |
| ExtendObsCounter | YES |
| Filename | C:\Users\Richard\AppData\Local\Temp\SAS Temporary Files\_TD6784_RICHARD-PC_\concord1.sas7bdat |

| Alphabetic List of Variables and Attributes | | | | |
|---|---|---|---|---|
| # | Variable | Type | Len | Label |
| 9 | ChPeople | Num | 8 | Change in number of people in household 1980 to 1981 |
| 6 | Education | Num | 8 | Years of Education of Head of Household |
| 1 | Household | Num | 8 | |
| 5 | Income | Num | 8 | |
| 10 | People80 | Num | 8 | Number of people in household in 1980 |
| 8 | People81 | Num | 8 | Number of people in household in 1981 |
| 7 | Retired | Num | 8 | Head of household is retired (1 = yes, 0 = no) |
| 4 | Water79 | Num | 8 | Household Water Consumption in 1979 (in cubic feet) |
| 3 | Water80 | Num | 8 | Household Water Consumption in 1980 (in cubic feet) |
| 2 | Water81 | Num | 8 | Household Water Consumption in 1981 (in cubic feet) |

# Histograms for 1981 Water Consumption

*The UNIVARIATE Procedure*
*Variable:   Water81   (Household Water Consumption in 1981 (in cubic feet))*

| Moments | | | |
|---|---:|---|---:|
| N | 496 | Sum Weights | 496 |
| Mean | 2298.3871 | Sum Observations | 1140000 |
| Std Deviation | 1486.1235 | Variance | 2208563.05 |
| Skewness | 1.72754384 | Kurtosis | 4.71525305 |
| Uncorrected SS | 3713400000 | Corrected SS | 1093238710 |
| Coeff Variation | 64.6594083 | Std Error Mean | 66.7289149 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| **Location** | | **Variability** | |
| Mean | 2298.387 | Std Deviation | 1486 |
| Median | 2050.000 | Variance | 2208563 |
| Mode | 1000.000 | Range | 10000 |
| | | Interquartile Range | 1700 |

| Tests for Location: Mu0=0 | | | | |
|---|---|---:|---|---|
| **Test** | **Statistic** | | **p Value** | |
| Student's t | t | 34.44365 | Pr > \|t\| | <.0001 |
| Sign | M | 248 | Pr >= \|M\| | <.0001 |
| Signed Rank | S | 61628 | Pr >= \|S\| | <.0001 |

| Tests for Normality | | | | |
|---|---|---:|---|---|
| **Test** | **Statistic** | | **p Value** | |
| Shapiro-Wilk | W | 0.874784 | Pr < W | <0.0001 |
| Kolmogorov-Smirnov | D | 0.111118 | Pr > D | <0.0100 |
| Cramer-von Mises | W-Sq | 1.803446 | Pr > W-Sq | <0.0050 |
| Anderson-Darling | A-Sq | 11.8592 | Pr > A-Sq | <0.0050 |

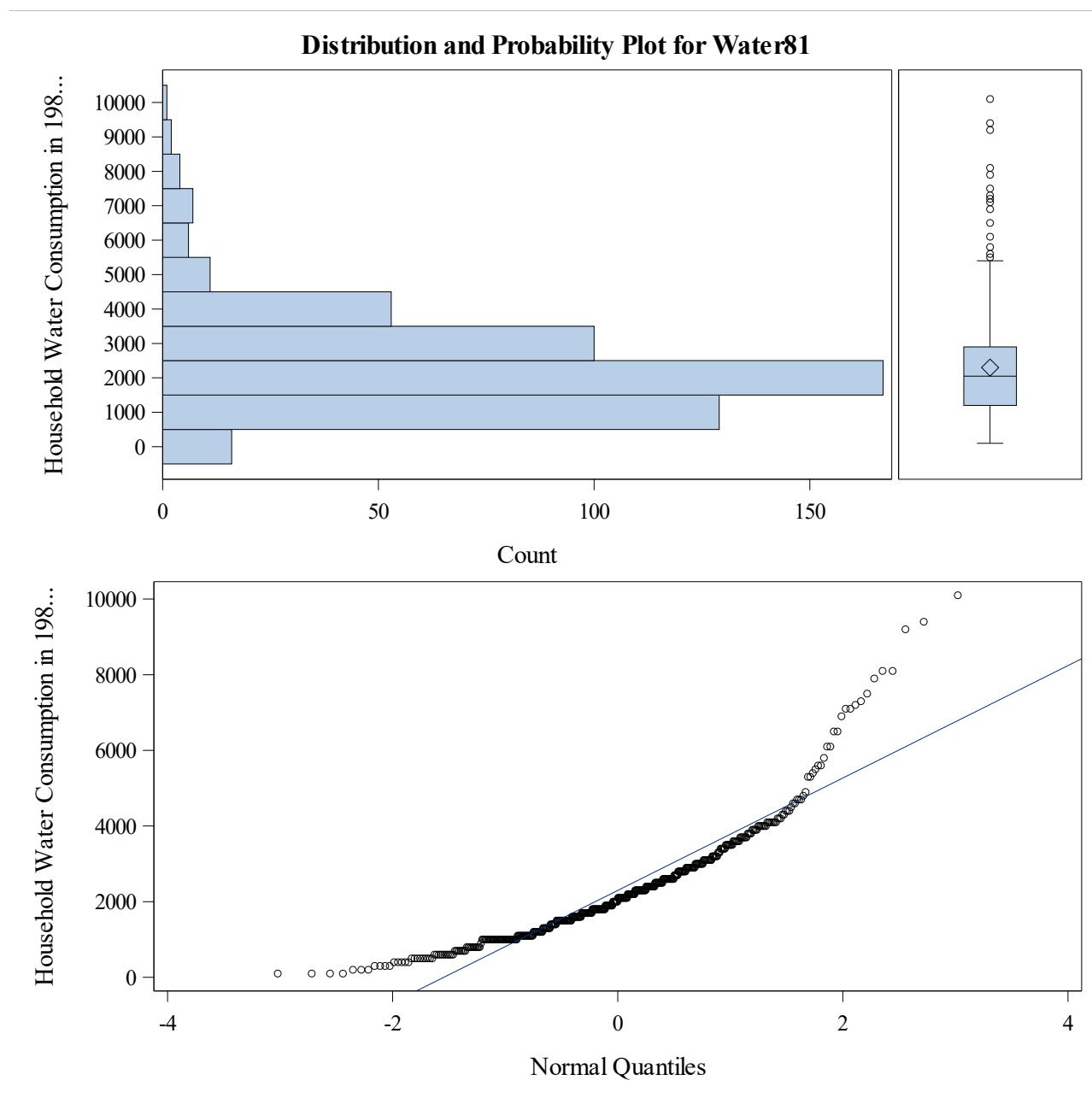**Histograms for 1981 Water Consumption**

*The UNIVARIATE Procedure*
  *Variable:   Water81   (Household Water Consumption in 1981 (in cubic feet))*

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 100% Max | 10100 |
| 99% | 8100 |
| 95% | 4800 |
| 90% | 4000 |
| 75% Q3 | 2900 |
| 50% Median | 2050 |
| 25% Q1 | 1200 |
| 10% | 800 |
| 5% | 500 |
| 1% | 200 |
| 0% Min | 100 |

| Extreme Observations | | | |
|---|---|---|---|
| Lowest | | Highest | |
| Value | Obs | Value | Obs |
| 100 | 330 | 8100 | 79 |
| 100 | 175 | 8100 | 451 |
| 100 | 105 | 9200 | 125 |
| 100 | 69 | 9400 | 62 |
| 200 | 396 | 10100 | 124 |

# Histograms for 1981 Water Consumption

## The UNIVARIATE Procedure

**Distribution and Probability Plot for Water81**

# Histograms for 1981 Water Consumption

## The UNIVARIATE Procedure



Distribution of Water81



Distribution of Water81
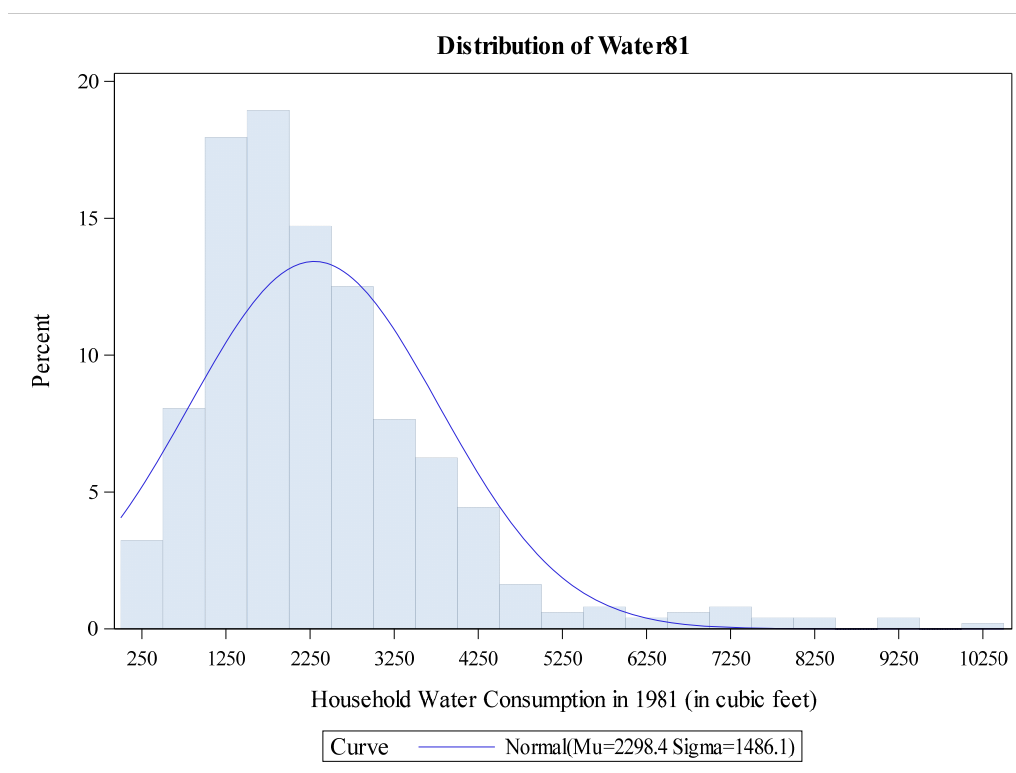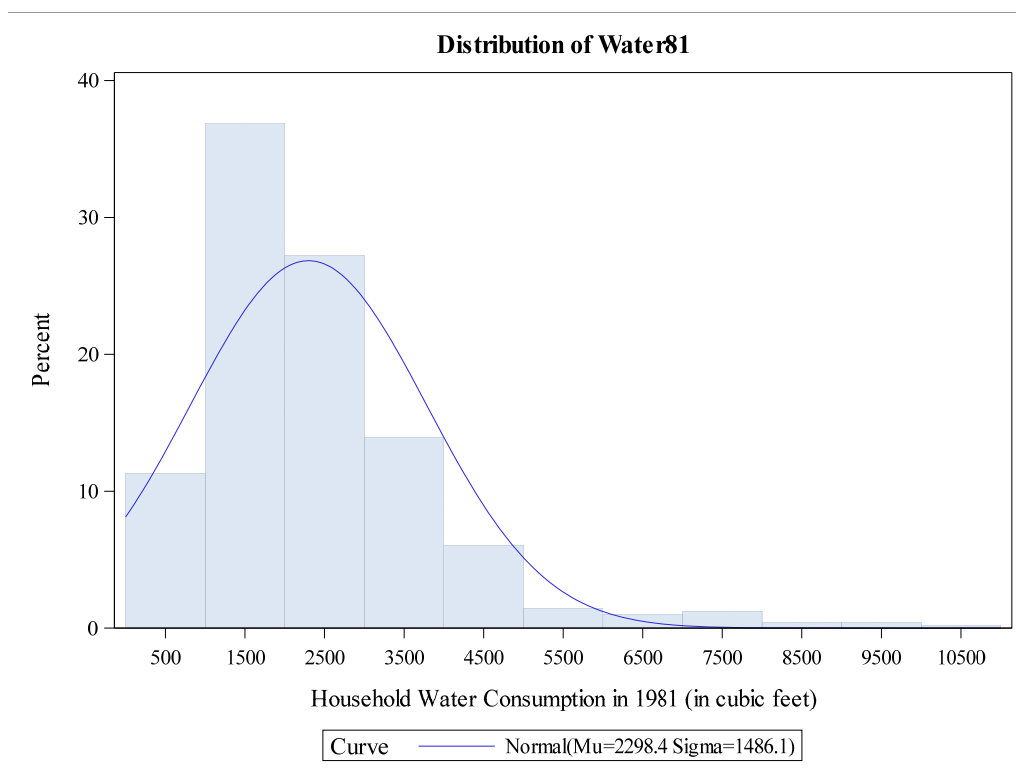
# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

*The UNIVARIATE Procedure*
*Variable:   srtWater81   (Square Root of Water Consumption in 1981)*

| Moments | | | |
|---|---|---|---|
| N | 496 | Sum Weights | 496 |
| Mean | 45.6617339 | Sum Observations | 22648.22 |
| Std Deviation | 14.6227307 | Variance | 213.824252 |
| Skewness | 0.51501016 | Kurtosis | 0.92187145 |
| Uncorrected SS | 1140000 | Corrected SS | 105843.005 |
| Coeff Variation | 32.0240372 | Std Error Mean | 0.65657999 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| Location | | Variability | |
| Mean | 45.66173 | Std Deviation | 14.62273 |
| Median | 45.27356 | Variance | 213.82425 |
| Mode | 31.62278 | Range | 90.49876 |
| | | Interquartile Range | 19.21063 |

| Tests for Location: Mu0=0 | | | | |
|---|---|---|---|---|
| Test | | Statistic | p Value | |
| Student's t | t | 69.54482 | Pr > \|t\| | <.0001 |
| Sign | M | 248 | Pr >= \|M\| | <.0001 |
| Signed Rank | S | 61628 | Pr >= \|S\| | <.0001 |

| Tests for Normality | | | | |
|---|---|---|---|---|
| Test | | Statistic | p Value | |
| Shapiro-Wilk | W | 0.980477 | Pr < W | <0.0001 |
| Kolmogorov-Smirnov | D | 0.055605 | Pr > D | <0.0100 |
| Cramer-von Mises | W-Sq | 0.221687 | Pr > W-Sq | <0.0050 |
| Anderson-Darling | A-Sq | 1.723555 | Pr > A-Sq | <0.0050 |

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 100% Max | 100.4988 |
| 99% | 90.0000 |
| 95% | 69.2820 |
| 90% | 63.2456 |
| 75% Q3 | 53.8516 |

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

*The UNIVARIATE Procedure*
*Variable:   srtWater81   (Square Root of Water Consumption in 1981)*

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 50% Median | 45.2736 |
| 25% Q1 | 34.6410 |
| 10% | 28.2843 |
| 5% | 22.3607 |
| 1% | 14.1421 |
| 0% Min | 10.0000 |
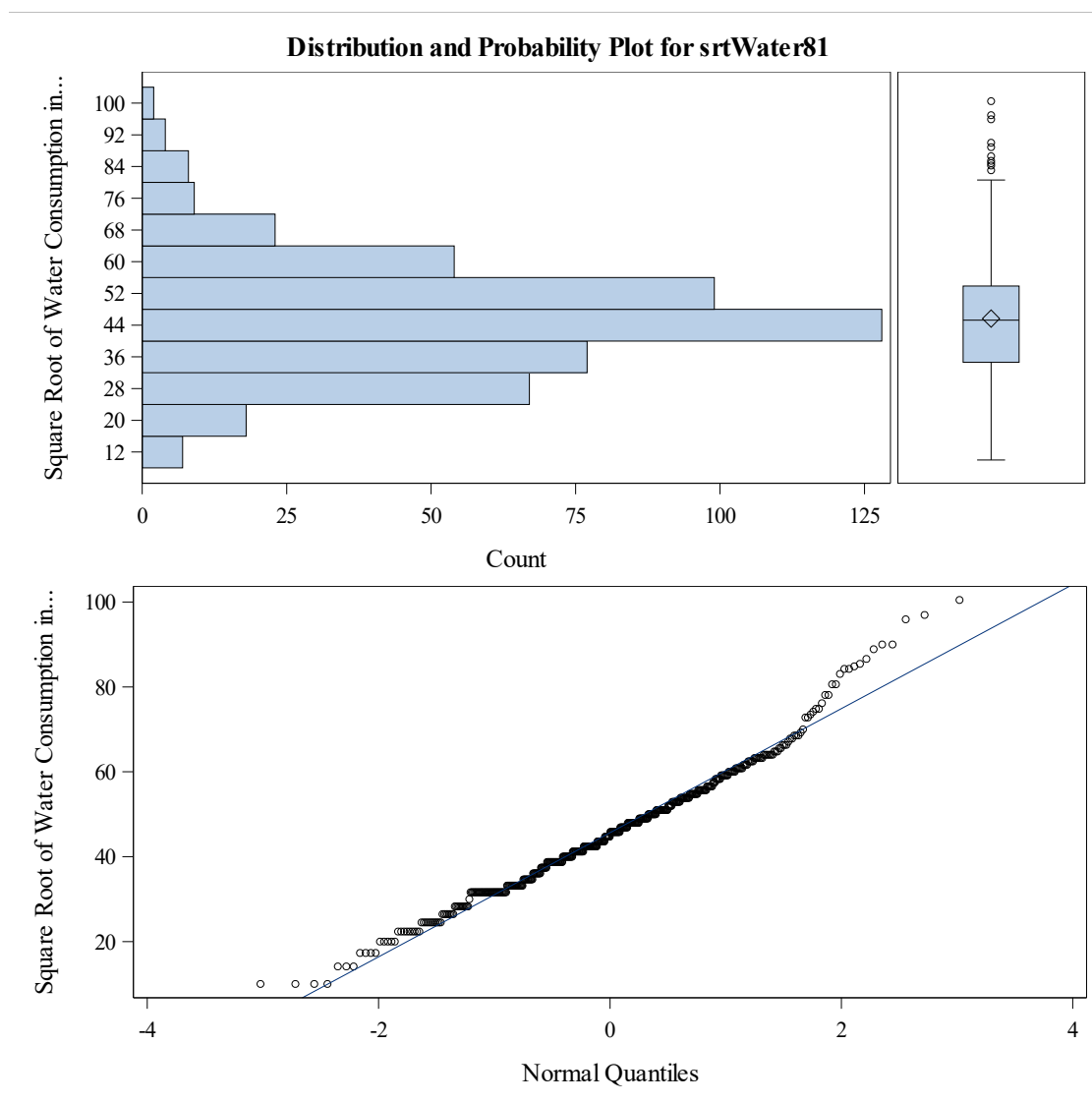
**Distribution and Probability Plot for srtWater81**

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

*The UNIVARIATE Procedure*
*Variable: crtWater81 (Cube Root of Water Consumption in 1981)*

| Moments | | | |
|---|---|---|---|
| N | 496 | Sum Weights | 496 |
| Mean | 12.626314 | Sum Observations | 6262.65175 |
| Std Deviation | 2.73864065 | Variance | 7.50015262 |
| Skewness | 0.09438616 | Kurtosis | 0.60484853 |
| Uncorrected SS | 82786.783 | Corrected SS | 3712.57554 |
| Coeff Variation | 21.6899457 | Std Error Mean | 0.1229686 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| Location | | Variability | |
| Mean | 12.62631 | Std Deviation | 2.73864 |
| Median | 12.70250 | Variance | 7.50015 |
| Mode | 10.00000 | Range | 16.97433 |
| | | Interquartile Range | 3.63385 |

| Tests for Location: Mu0=0 | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Student's t | t | 102.6792 | Pr > \|t\| | <.0001 |
| Sign | M | 248 | Pr >= \|M\| | <.0001 |
| Signed Rank | S | 61628 | Pr >= \|S\| | <.0001 |

| Tests for Normality | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Shapiro-Wilk | W | 0.991332 | Pr < W | 0.0053 |
| Kolmogorov-Smirnov | D | 0.05588 | Pr > D | <0.0100 |
| Cramer-von Mises | W-Sq | 0.129674 | Pr > W-Sq | 0.0457 |
| Anderson-Darling | A-Sq | 1.004639 | Pr > A-Sq | 0.0127 |

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 100% Max | 21.61592 |
| 99% | 20.08299 |
| 95% | 16.86865 |
| 90% | 15.87401 |
| 75% Q3 | 14.26043 |

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

### The UNIVARIATE Procedure
### Variable:  crtWater81  (Cube Root of Water Consumption in 1981)

| Quantiles (Definition 5) | |
|---|---|
| **Level** | **Quantile** |
| **50% Median** | 12.70250 |
| **25% Q1** | 10.62659 |
| **10%** | 9.28318 |
| **5%** | 7.93701 |
| **1%** | 5.84804 |
| **0% Min** | 4.64159 |
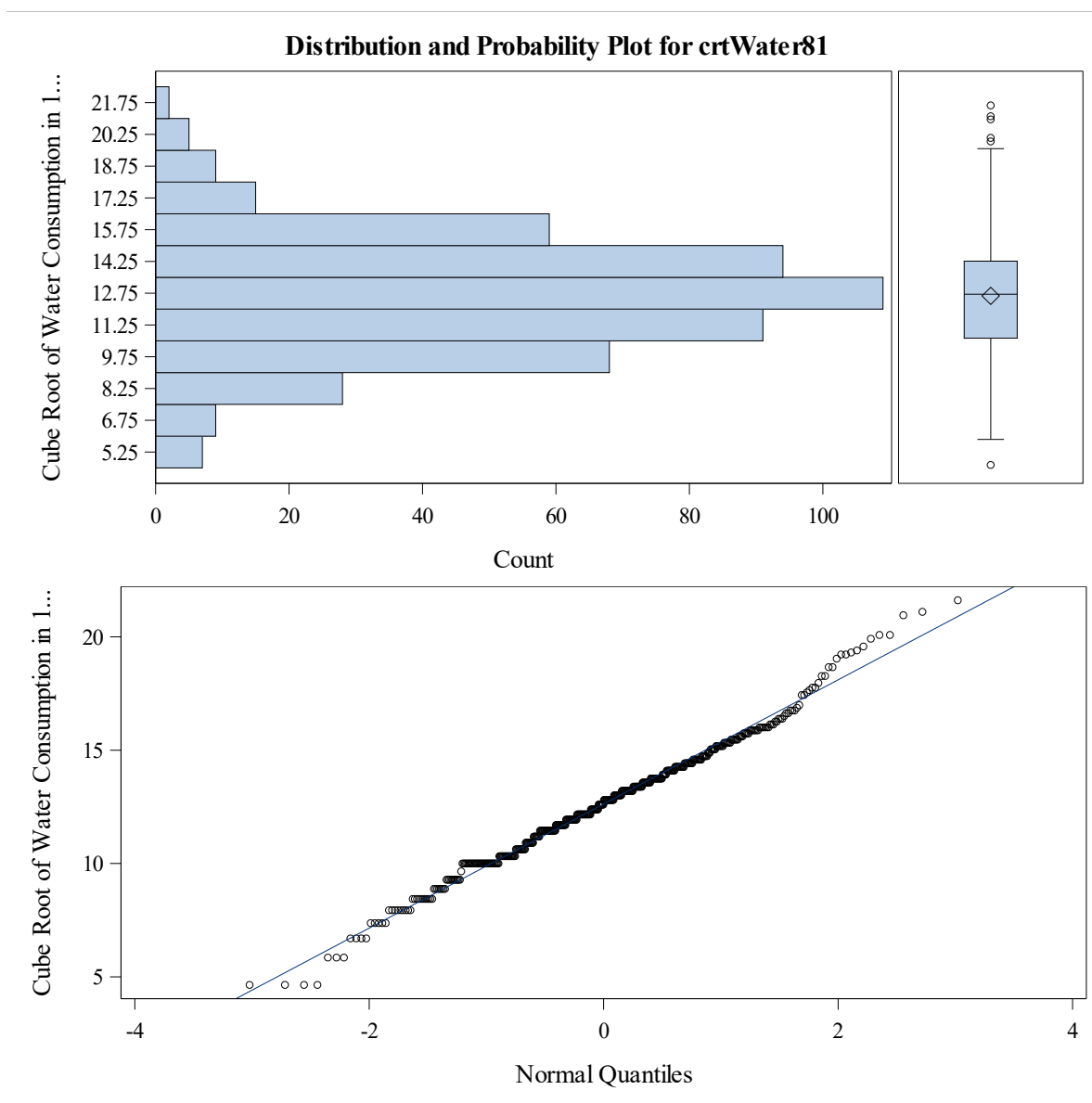


Distribution and Probability Plot for crtWater81

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

## The UNIVARIATE Procedure



Distribution of crtWater81

Curve — Normal(Mu=12.626 Sigma=2.7386)
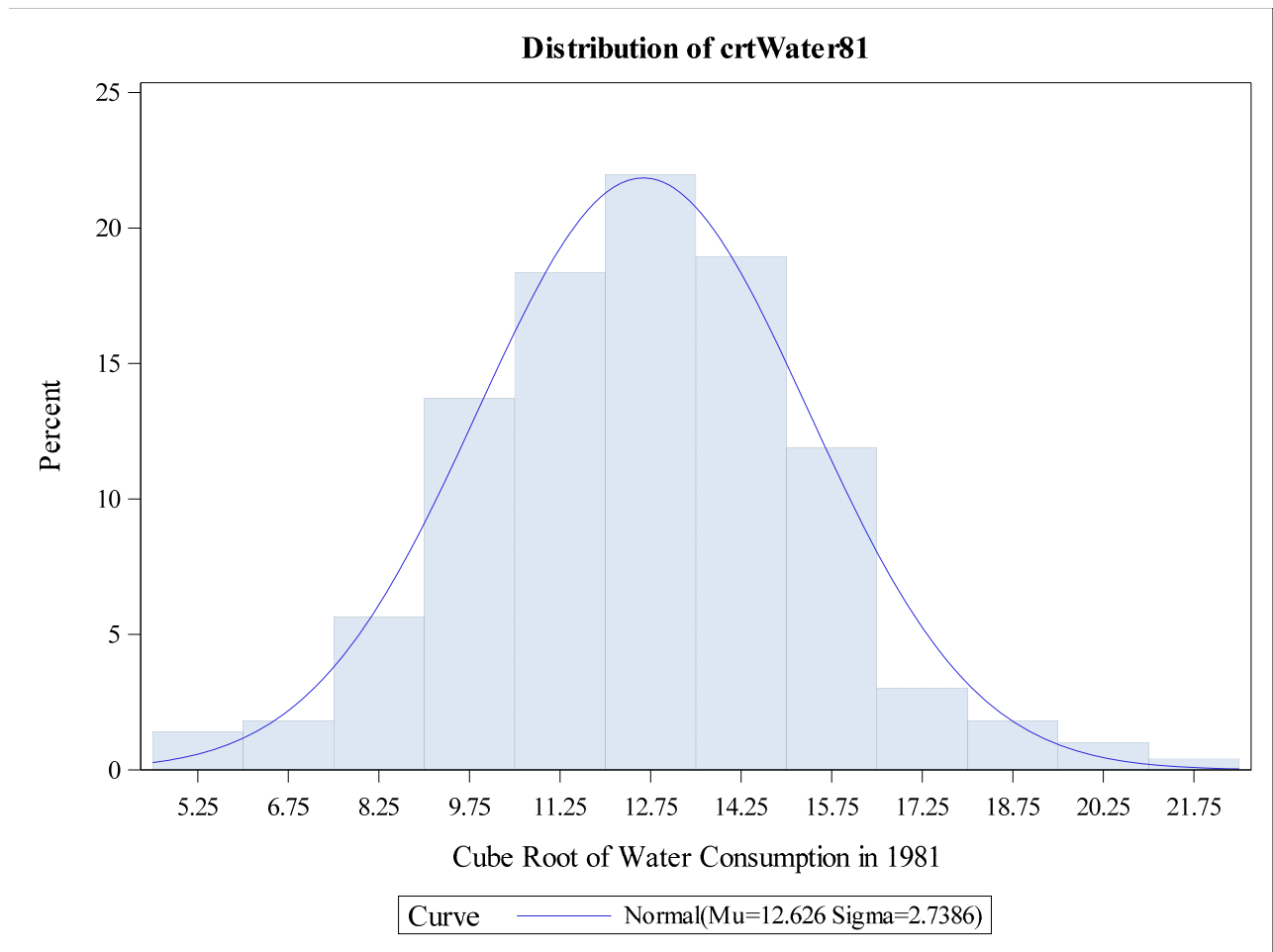
# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

### The UNIVARIATE Procedure
### Variable:   logWater81   (Log of Water Consumption in 1981)

| Moments | | | |
|---|---:|---|---:|
| N | 496 | Sum Weights | 496 |
| Mean | 7.53096199 | Sum Observations | 3735.35715 |
| Std Deviation | 0.69926608 | Variance | 0.48897305 |
| Skewness | -0.8984223 | Kurtosis | 2.1625174 |
| Uncorrected SS | 28372.8743 | Corrected SS | 242.041661 |
| Coeff Variation | 9.28521592 | Std Error Mean | 0.03139797 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| **Location** | | **Variability** | |
| Mean | 7.530962 | Std Deviation | 0.69927 |
| Median | 7.625298 | Variance | 0.48897 |
| Mode | 6.907755 | Range | 4.61512 |
| | | Interquartile Range | 0.88239 |

| Tests for Location: Mu0=0 | | | | |
|---|---|---:|---|---|
| **Test** | | **Statistic** | **p Value** | |
| Student's t | t | 239.855 | Pr > \|t\| | <.0001 |
| Sign | M | 248 | Pr >= \|M\| | <.0001 |
| Signed Rank | S | 61628 | Pr >= \|S\| | <.0001 |

| Tests for Normality | | | | |
|---|---|---:|---|---|
| **Test** | | **Statistic** | **p Value** | |
| Shapiro-Wilk | W | 0.955801 | Pr < W | <0.0001 |
| Kolmogorov-Smirnov | D | 0.085415 | Pr > D | <0.0100 |
| Cramer-von Mises | W-Sq | 0.67678 | Pr > W-Sq | <0.0050 |
| Anderson-Darling | A-Sq | 4.189554 | Pr > A-Sq | <0.0050 |

| Quantiles (Definition 5) | |
|---|---:|
| **Level** | **Quantile** |
| 100% Max | 9.22029 |
| 99% | 8.99962 |
| 95% | 8.47637 |

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

### The UNIVARIATE Procedure
### Variable:   logWater81   (Log of Water Consumption in 1981)

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 90% | 8.29405 |
| 75% Q3 | 7.97247 |
| 50% Median | 7.62530 |
| 25% Q1 | 7.09008 |
| 10% | 6.68461 |
| 5% | 6.21461 |
| 1% | 5.29832 |
| 0% Min | 4.60517 |

| Extreme Observations | | | |
|---|---|---|---|
| Lowest | | Highest | |
| Value | Obs | Value | Obs |
| 4.60517 | 330 | 8.99962 | 79 |
| 4.60517 | 175 | 8.99962 | 451 |
| 4.60517 | 105 | 9.12696 | 125 |
| 4.60517 | 69 | 9.14846 | 62 |
| 5.29832 | 396 | 9.22029 | 124 |

# SUMMARY STATISTICS AND PLOTS FOR TRANSFORMED 1981 WATER CONSUMPTION

## *The UNIVARIATE Procedure*

**Distribution and Probability Plot for logWater81**
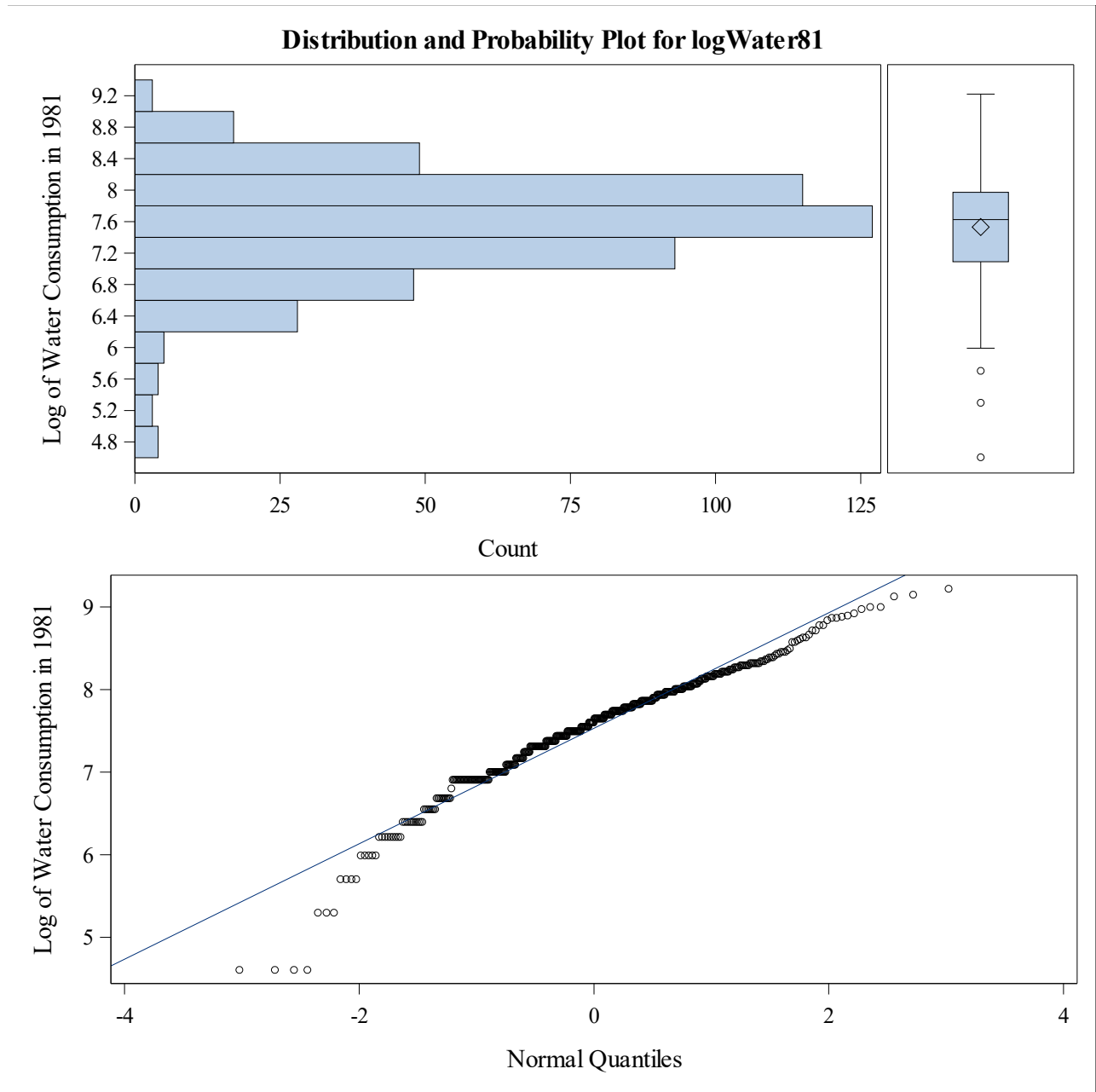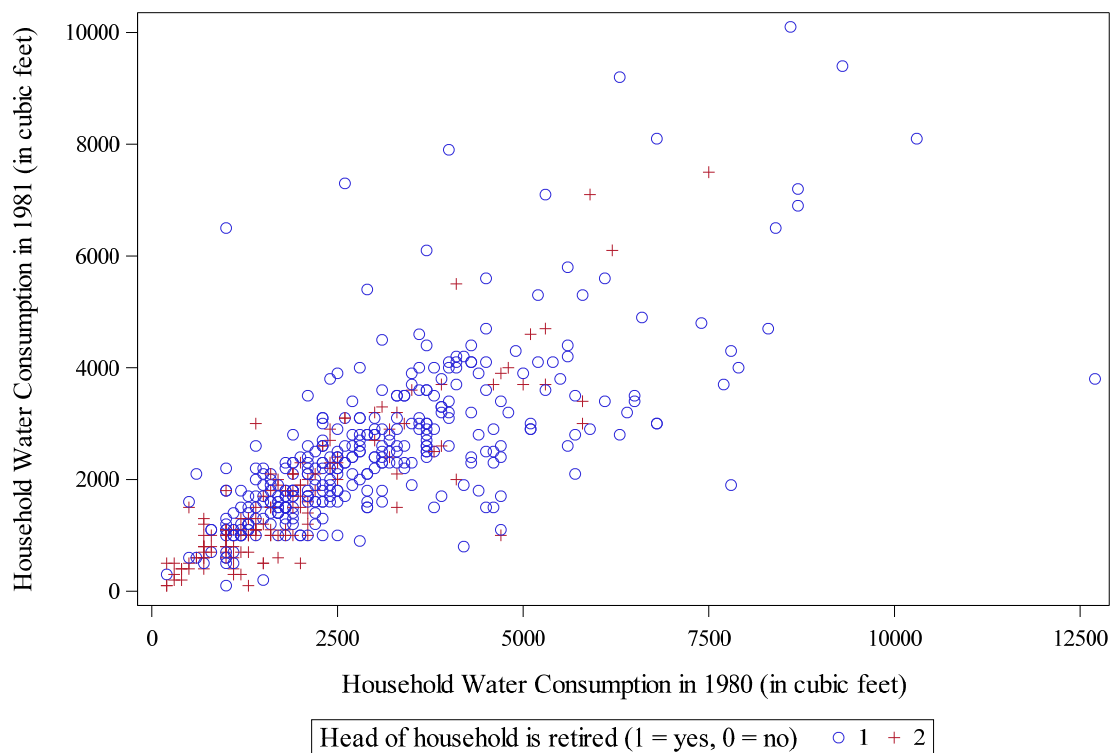
**SCATTER PLOT OF 1981 WATER CONSUMPTION AGAINST 1980 WATER CONSUMPTION**
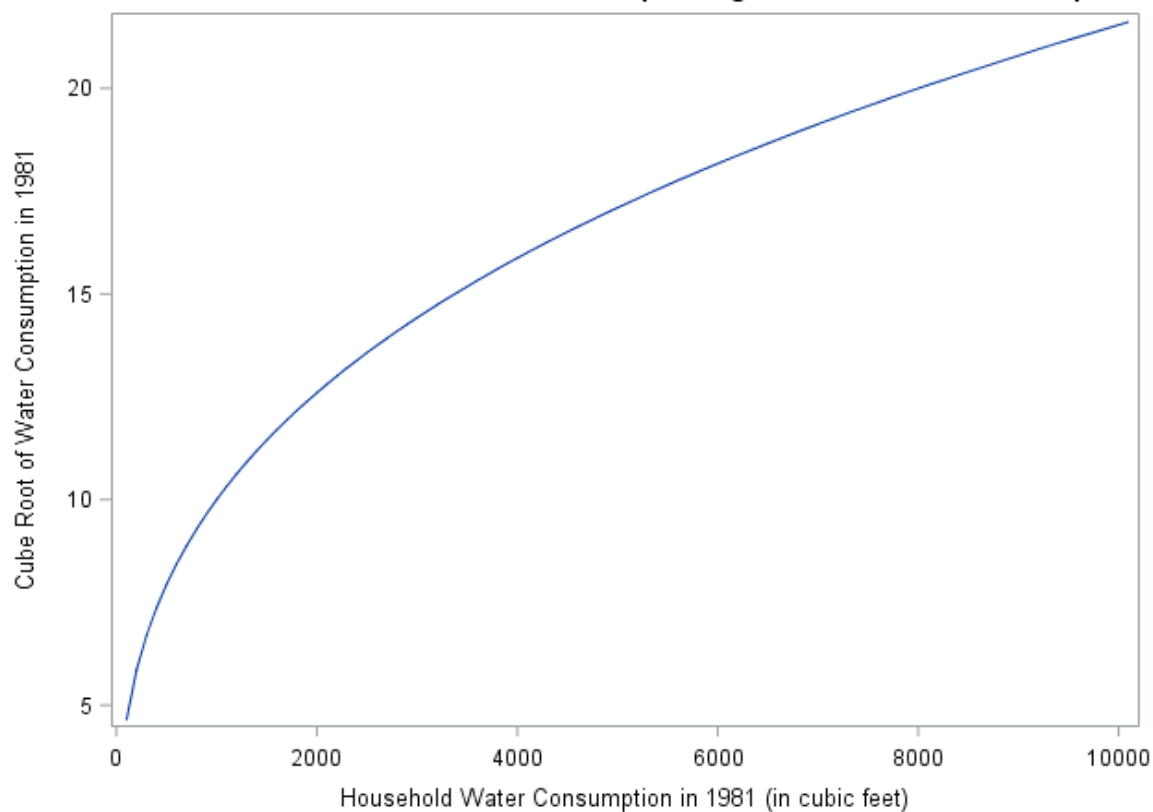
## *Correlations among Variable in Concord Water Data*

### *The CORR Procedure*

| 9 Variables: | Water81 Water80 Water79 Income Education<br>Retired People81 ChPeople People80 |
|---|---|

| Simple Statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **N** | **Mean** | **Std Dev** | **Sum** | **Minimum** | **Maximum** | **Label** |
| **Water81** | 496 | 2298 | 1486 | 1140000 | 100.00000 | 10100 | Household Water Consumption in 1981 (in cubic feet) |
| **Water80** | 496 | 2732 | 1764 | 1355100 | 200.00000 | 12700 | Household Water Consumption in 1980 (in cubic feet) |
| **Water79** | 496 | 2692 | 2003 | 1335353 | -1.00000 | 14500 | Household Water Consumption in 1979 (in cubic feet) |
| **Income** | 496 | 23077 | 13058 | 11446000 | 2000 | 100000 | |
| **Education** | 496 | 14.00403 | 3.09055 | 6946 | 6.00000 | 20.00000 | Years of Education of Head of Household |
| **Retired** | 496 | 1.29435 | 0.45621 | 642.00000 | 1.00000 | 2.00000 | Head of household is retired (1 = yes, 0 = no) |
| **People81** | 496 | 3.07258 | 1.65718 | 1524 | 1.00000 | 10.00000 | Number of people in household in 1981 |
| **ChPeople** | 496 | -0.03831 | 0.48466 | -19.00000 | -3.00000 | 3.00000 | Change in number of people in household 1980 to 1981 |
| **People80** | 496 | 3.11089 | 1.65810 | 1543 | 1.00000 | 10.00000 | Number of people in household in 1980 |

# *Correlations among Variable in Concord Water Data*

## *The CORR Procedure*

| Pearson Correlation Coefficients, N = 496<br>Prob > \|r\| under H0: Rho=0 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Water81** | **Water80** | **Water79** | **Income** | **Education** | **Retired** | **People81** | **ChPeople** | **People80** |
| **Water81**<br>Household Water Consumption in 1981 (in cubic feet) | 1.00000 | 0.76479<br><.0001 | 0.63979<br><.0001 | 0.41779<br><.0001 | 0.04038<br>0.3695 | -0.27313<br><.0001 | 0.61831<br><.0001 | 0.06611<br>0.1415 | 0.59864<br><.0001 |
| **Water80**<br>Household Water Consumption in 1980 (in cubic feet) | 0.76479<br><.0001 | 1.00000 | 0.60091<br><.0001 | 0.33705<br><.0001 | 0.09822<br>0.0287 | -0.29193<br><.0001 | 0.52510<br><.0001 | -0.03117<br>0.4885 | 0.53392<br><.0001 |
| **Water79**<br>Household Water Consumption in 1979 (in cubic feet) | 0.63979<br><.0001 | 0.60091<br><.0001 | 1.00000 | 0.29560<br><.0001 | 0.05948<br>0.1860 | -0.20176<br><.0001 | 0.44993<br><.0001 | 0.00136<br>0.9758 | 0.44928<br><.0001 |
| **Income** | 0.41779<br><.0001 | 0.33705<br><.0001 | 0.29560<br><.0001 | 1.00000 | 0.34625<br><.0001 | -0.38056<br><.0001 | 0.31128<br><.0001 | 0.09112<br>0.0425 | 0.28447<br><.0001 |
| **Education**<br>Years of Education of Head of Household | 0.04038<br>0.3695 | 0.09822<br>0.0287 | 0.05948<br>0.1860 | 0.34625<br><.0001 | 1.00000 | -0.17421<br><.0001 | 0.05872<br>0.1917 | 0.00550<br>0.9028 | 0.05708<br>0.2045 |
| **Retired**<br>Head of household is retired (1 = yes, 0 = no) | -0.27313<br><.0001 | -0.29193<br><.0001 | -0.20176<br><.0001 | -0.38056<br><.0001 | -0.17421<br><.0001 | 1.00000 | -0.37569<br><.0001 | -0.05854<br>0.1930 | -0.35837<br><.0001 |
| **People81**<br>Number of people in household in 1981 | 0.61831<br><.0001 | 0.52510<br><.0001 | 0.44993<br><.0001 | 0.31128<br><.0001 | 0.05872<br>0.1917 | -0.37569<br><.0001 | 1.00000 | 0.14433<br>0.0013 | 0.95726<br><.0001 |
| **ChPeople**<br>Change in number of people in household 1980 to 1981 | 0.06611<br>0.1415 | -0.03117<br>0.4885 | 0.00136<br>0.9758 | 0.09112<br>0.0425 | 0.00550<br>0.9028 | -0.05854<br>0.1930 | 0.14433<br>0.0013 | 1.00000 | -0.14805<br>0.0009 |
| **People80**<br>Number of people in household in 1980 | 0.59864<br><.0001 | 0.53392<br><.0001 | 0.44928<br><.0001 | 0.28447<br><.0001 | 0.05708<br>0.2045 | -0.35837<br><.0001 | 0.95726<br><.0001 | -0.14805<br>0.0009 | 1.00000 |

# Correlations among Variable in Concord Water Data

## The CORR Procedure

### Scatter Plot Matrix