```
* Programmed by Charles Hwang *
* Coded in SAS OnDemand       *
* Monday, March 25, 2019      *
* Course: STAT 303            *
* Title: Homework 4           *;
/* 1 */ Data Arsenic;
Input Arsenic;
Datalines;
9.722
10.162
9.976
9.787
9.474
10.113
10.157
9.556
9.667
9.809
10.424
9.288
;
Run;

/* 1a */
* H0: μ = 10 ppb of arsenic
  HA: μ < 10 ppb of arsenic
  α = .01 ;

/* 1b */ Proc ttest h0=10 alpha=0.01;
Var Arsenic;
Run;

/* 1c */ * t = -1.63, p = .1321/2 = .06605
  We fail to reject H0 at α = .01. There is insufficient evidence that the water contains
  less than 10 ppb of arsenic.;

/* 1d */ * The citizens should refrain from drinking the water. The hypothesis test fails to prove
  that the level of arsenic in the water is safe for human consumption. ;

/* 1e */ * 99% Confidence Interval: (9.5478, 10.1413)
  We are 99 percent confident that the true population mean of arsenic in the water is
  between 9.5478 and 10.1413 ppb. ;

/* 2a */ Data MSCE;
Infile "/home/chwang10/sasuser.v94/MSCE.txt" DLM='09'X;
Length Race$ 9;
Input Race$ MSCE;
If Race="Caucasian" OR Race="NativeAM";
Run;

/* 2b */ ODS graphics on;
Proc ttest Plots=Histogram;
Class Race;
Var MSCE;
Run;
/* 2b(a) */
* H0: The mean MSCE values of Caucasians and Native Americans are the same
  HA: The mean MSCE values of Caucasians and Native Americans are not the same
  α = .05 ;
/* 2b(b) */ * We should use pooled variances because the Equality-of-Variances Test was not
  significant (F = 1.30, p = .6946). ;
/* 2b(c) */ * t = -2.15 ;
```

```
/* 2b(d) */ * p = .0471
  We reject H0 at α = .05. There is sufficient evidence that the mean MSCE values of Caucasians and
  Native Americans are different. ;
/* 2b(e) */ * 95% Confidence Interval: (-0.9863, -0.00719)
  We are 95 percent confident that the mean MSCE values of Caucasians minus the mean MSCE values of
  Native Americans is between -0.9863 and -0.00719. ;

/* 2c */ Proc npar1way wilcoxon;
Class Race;
Var MSCE;
Run;
/* 2c(a) */
* H0: The distribution of MSCE values between Caucasians and Native Americans are the same
  HA: The distribution of MSCE values between Caucasians and Native Americans are not the same
  α = .05 ;
/* 2c(b) */ * S = 99.00 ;
/* 2c(c) */ * p = .0617
  We fail to reject reject H0 at α = .05. There is insufficient evidence that the distribution of
  MSCE values between Caucasians and Native Americans are different. ;

/* 2d */ * There is a clear pattern in the QQ plot for Caucasians that suggests the distribution is
  not normal. A similar but less distinct pattern exists in the QQ plot for Native Americans. This
  points to the Wilcoxon Sum-Rank Test being the better method because it does not require the
  assumption of normal distributions. In this problem, it is essential to choose the correct test
  because the conclusions of the two tests were different at the same α-level. ;

/* 3a */ Data Score;
Infile "/home/chwang10/sasuser.v94/readingscores.txt";
Input Pre Post;
Diff=Post-Pre;
Run;

/* 3b */ ODS graphics on;
Proc ttest Plots=Histogram h0=10 side=u;
Paired Post*Pre;
Run;
/* 3b(a) */
* H0: The mean test score is the same before and after the reading program
  HA: The mean test score is less before the reading program than after
  α = .05 ;
/* 3b(b) */ * t = 1.91 ;
/* 3b(c) */ * p = .0386
  We reject H0 at α = .05. There is sufficient evidence that the mean test score before and after
  the reading program are different. ;
/* 3b(d) */ * 95% Confidence Interval: (10.3054, ∞)
  We are 95 percent confident that the mean post-test score minus the mean pre-test score is
  greater than 10.3054. ;

/* 3c */ Ods select TestsForLocation;
Proc Univariate mu0=10;
Var Diff;
Run;
/* 3c(a) */
* H0: The median test score is the same before and after the reading program
  HA: The median test score is less before the reading program than after
  α = .05 ;
/* 3c(b) */ * S = 24 ;
/* 3c(c) */ * p = .1361/2 = .06805
  We fail to reject H0 at α = .05. There is insufficient evidence that the median test score is less
  before the reading program than after.
  This p-value is greater than the p-value obtained from the t-test (.06805 > .03865). ;

/* 3d */ * The histogram for the distribution of the difference appears to be right skewed and
```

may not be normal enough to apply the paired t-test on. This points to the Wilcoxon Sum-Rank Test being the better method because it does not require the assumption of normal distributions. In this problem, it is essential to choose the correct test because the conclusions of the two tests were different at the same α-level. ;