

Q3.4

1. Metric scale. In particular, the t-test is not valid for variables which only have an ordinal structure.
2. IID data. In particular, each pair of measurements (X_i , Y_i) is drawn from the same distribution, independently of all other pairs.
3. The distribution of the difference between measurements has no major deviations from normality, considering the sample size. In particular, the t-test is invalid for highly skewed distributions when sample size is larger than 30. It may also be invalid for very highly skewed distributions at higher sample sizes.

```
##           X           year           id           prottemp
## Min.      : 1.0    Min.      :2004    Min.      : 4.0    Min.      : 0.00
## 1st Qu.:201.2    1st Qu.:2004    1st Qu.: 728.8    1st Qu.: 50.00
## Median :401.5    Median :2004    Median :1373.5    Median : 60.00
## Mean   :401.5    Mean   :2004    Mean   :1381.9    Mean   : 65.56
## 3rd Qu.:601.8    3rd Qu.:2004    3rd Qu.:2053.5    3rd Qu.: 85.00
## Max.    :802.0    Max.    :2004    Max.    :2808.0    Max.    :100.00
##      cathtemp
## Min.      : 0.00
## 1st Qu.: 50.00
## Median : 60.00
## Mean   : 63.16
## 3rd Qu.: 85.00
## Max.    :100.00

##   X year id prottemp cathtemp
## 1 1 2004 4      50      50
## 2 2 2004 7       5      85
## 3 3 2004 9      50      50
## 4 4 2004 14     60      60
## 5 5 2004 21    100       0
## 6 6 2004 24      5       5
```

1. Metric scale: Based on the way the thermometer was described to the participants, we could not say that the thermometer can be classified as metric. This data is not measuring any objective quantity (such as the number of deaths in 3.3). This is essentially a Likert scale but with more “levels”, making this ordinal data. For example, if someone gave a 50 rating for Protestants and someone else gave a 100 rating on Protestants, we cannot definitively say that the one who gave the 100 rating is twice as warm towards Protestants than the first person, or that we can add/subtract these feelings from each other in a meaningful way.
2. IID data: If the experimental design was done properly, the data can likely be considered independent, as survey-takers would not know each other and not have influence on one another. Assuming the GSS data is drawn from the same distribution, then the identically distributed assumption also holds.
3. The distribution of the difference between measurements has no major deviations from normality, considering the sample size: Since the data is not ordinal, not metric, assessment for normality does not apply. If we were working with metric data that wasn’t very skewed, the sample size is sufficiently large (802) and would allow for the CLT to be applied.

Conclusion: The data is likely IID. However, the data is ordinal and not metric and therefore can’t be considered normal. This makes the paired t-test an inappropriate choice for this dataset.