Zander Gonzales

Statistical Analysis Project

ME 275

Dr. Steven Charles

# Statistical Analysis of Punxsutawney Phil and Following Weather

**Abstract**

Every year on February 2nd Punxsutawney Phil predicts whether there will continue to be winter or whether there will be an early spring. This report takes his predictions as well as weather data from NOAA's National Climatic Data Center, and performs a double sample hypothesis mean test to find if Punxsutawney Phil's predictions are statistically significant. The data includes 115 years for which Punxsutawney Phil's prediction was recorded and temperatures were also recorded. 100 of these predictions were for winter, and 15 were for spring. The p value for his winter predictions was 0.2906, while the p value for his spring predictions was 0.0092. Thus we conclude that it is not statistically significant if Punxsutawney predicts a continued winter, but it is significant if he predicts an early spring.

## Introduction

Every year on February 2nd Punxsutawney Phil, a groundhog, comes out and looks around to see if he can see his shadow. If he sees his shadow, the prediction is that there will be another six weeks of winter that year; if he doesn't see his shadow, the prediction is that there will be an early spring that year. While this has been a tradition in the United States going back to 1887, the objective of this work is to see if there is any true correlation between Punxsutawney Phil seeing his shadow and the weather following that event. This will allow us to see if it is statistically significant whether Punxsutawney Phil sees his shadow or not, and what we might be able to expect from the weather based on that data.

## Methods

To find the information regarding whether Punxsutawney saw his shadow or not, I used data collected by the Punxsutawney Groundhog Club, which has tracked whether Punxsutawney Phil sees his shadow every year since 1887. Since we can't actually speak to Punxsutawney Phil, whether he "sees" his shadow is determined by a group called "The Inner Circle", who base it off of whether it is generally sunny or overcast on that day. There are some years early on where there is a recorded "No Record" and there is also one "Partial Shadow" recorded, so these were removed from the dataset.

The information about weather comes from NOAA's National Climatic Data Center, where they measure temperature in different regions of the United States and average them together per month. They get these temperature measurements from thousands of sources, including weather stations, buoys, and satellites. This raw data is then taken and processed to remove any errors or duplicate readings, and then published separated by region. This started in 1895, so any years before that were removed from the dataset.

To create the inferential statistic, there were two null hypotheses formed, based on whether Punxsutawney Phil saw his shadow or not. **If he saw his shadow, the null hypothesis** was that the mean temperature of the following month would be equal to or greater than the mean temperature of every year combined for the month. **The alternative** was that the mean would be lower than the total mean of all the data. **If he didn't see his shadow, the null hypothesis** was that the mean temperature of the following month would be equal to or less than the mean temperature of every year combined for the month. **The alternative** was that the mean would be greater than the total mean of all the data. This was a way of measuring whether it really would continue to be winter if he saw his shadow, and if there really would be an early spring if he didn't see his shadow. From there the hypotheses were tested using a double sample hypothesis test in Python, using the Scipy package (scipy.stats.ttest_ind()).

**Results**

Firstly, an analysis of the data that would be worked with was performed. As can be seen in Figure 1 below, the data from all the years combined matches very closely with the data for when Punxsutawney Phil did see his shadow. However, it is significantly different from the data for when he didn't see his shadow.

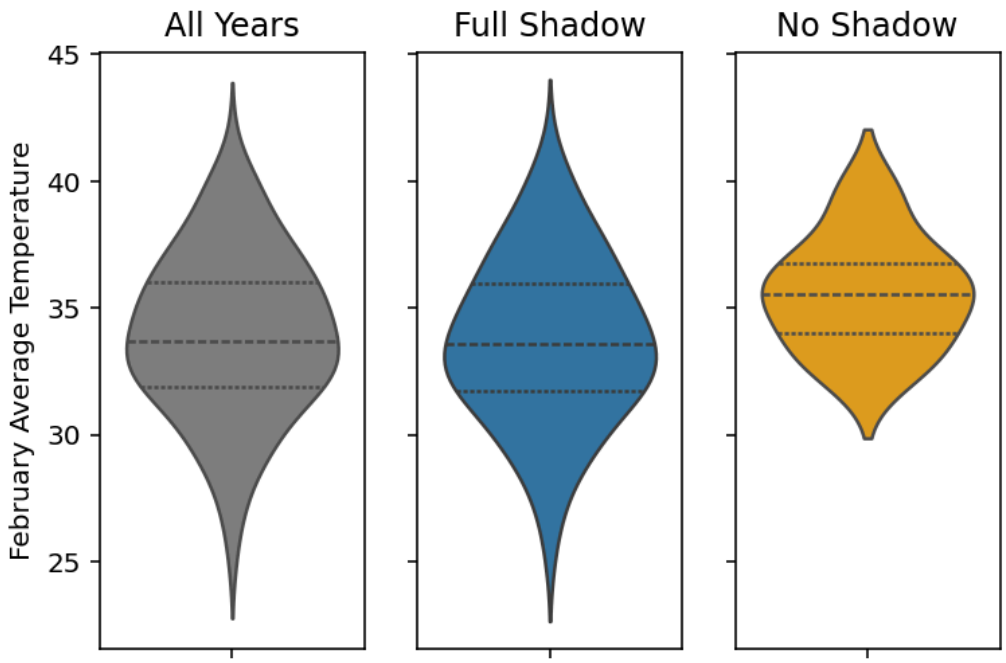


*Figure 1. A violin plot showing the average temperature in February, separated by whether Punxsutawney saw his shadow or not. The inside is divided into quartiles.*

This in part could be due to the fact that there are 115 data points total, and 100 of those include Punxsutawney Phil seeing his shadow while only 15 have him not seeing his shadow. This is more clearly shown in Table 1, where the numerical values are given.

*Table 1. Shows the descriptive statistics surrounding average temperature in February, separated by whether Punxsutawney Phil saw his shadow or not.*

| | Sample Size (Years) | Mean (℃) | Standard Deviation (℃) |
|---|---|---|---|
| All Years | 115 | 33.956 | 3.187 |
| Full Shadow | 100 | 33.713 | 3.249 |
| No Shadow | 15 | 35.578 | 2.189 |

The double sample hypothesis tests were run using Scipy in Python (scipy.stats.ttest_ind()), and the tests were run not assuming equal variance. The results of the tests are found below.

*Table 2. Contains the resultant p-value for a double sample hypothesis test, with the null hypotheses being that the mean temperature if he saw his shadow would be greater than or equal to the total mean temperature, and that the mean temperature if he didn't see his shadow would be less than or equal to the total mean temperature. The confidence intervals provided are a 95% confidence of the difference in means between the two samples.*

| | P-Value | Confidence Interval |
|---|---|---|
| Full Shadow | 0.2906 | [-∞, 0.4842] |
| No Shadow | 0.0092 | [0.5269, ∞] |

**Discussion**

In the context of the original question, the p values show an interesting significance. If Punxsutawney Phil doesn't see his shadow, thereby predicting a continued winter, there is almost a 30% chance the average temperature of the next month will be greater than the total average, putting it well above a 5% significance level. However, if Punxsutawney Phil sees his shadow and predicts an early spring, there is only a 0.92% chance that the average temperature of the next month will be less than the total average, meaning it is statistically significant and we can reject the null hypothesis. This means we accept the alternative

hypothesis that the average temperature of the following month will be greater than the total average temperature, in other words, an early spring.

The confidence intervals show a similar result. Because these are one sided tests it means that infinity will be included in them, but the first test includes 0 in it, while the second doesn’t. This means that the difference between Februarys where Punxsutawney Phil sees his shadow and all Februarys has a chance in the 95% confidence range of being 0. But if he doesn't see his shadow there, we can say with 95% confidence that there will be a difference in the mean temperature in the following February. Thus, we conclude that it is not statistically significant if Punxsutawney predicts a continued winter ($p= 0.2906$), but it is significant if he predicts an early spring ($p= 0.0092$).

**Statement on the use of artificial intelligence**

I did not use artificial intelligence for any calculations or data analysis.