An analysis on the temperature increase in the past four Summers at Hofstra
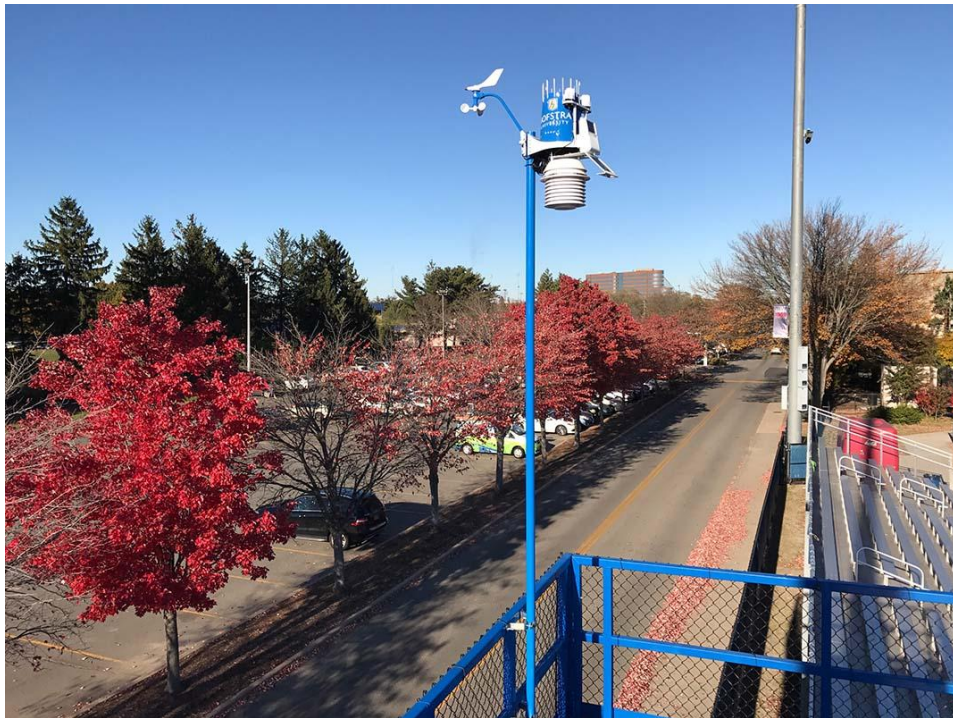
Wenxin Liu
MATH 138 Spring 2021
Hofstra University

Purpose/ Background

- Global warming
- Summers getting hotter
- POLAR BEARS!

From https://www.wwf.org.uk/learn/wildlife/polar-bears

```
Record ID,Timestamp,Thermometer
589392,2017-06-01 00:00:11,62.6
589395,2017-06-01 00:02:11,62.6
589398,2017-06-01 00:03:11,62.7
589399,2017-06-01 00:04:11,62.7
589402,2017-06-01 00:05:12,62.7
589405,2017-06-01 00:06:12,62.8
589407,2017-06-01 00:07:13,62.9
589408,2017-06-01 00:09:13,63.0
```

From nassau-ny.weatherstem.com/hofstrasoccer

Methods

- ANOVA TEST ON 4 YEARS

Values	
avg_17	73.0223684210526
avg_18	74.9224705882353
avg_19	75.2433720930233
avg_20	75.884512195122
sd_17	5.4846330459861
sd_18	5.83377622248279
sd_19	5.69336881777494
sd_20	5.18358023764428

Files   Plots   Packages   Help   Viewer

- 3 PAIRWISE HYPOTHESIS TESTINGS FOR TWO MEANS ON TWO CONSECUTIVE YEARS

ANOVA test

H0: four means are the same

H1: at least one of the four means is not the same

```
> June.Aug17.20 <- read.csv("~/R/June-Aug17-20.txt")
>   View(June.Aug17.20)
> tada = aov(June.Aug17.20$da_TEMP~June.Aug17.20$da_YR)
> summary(tada)
                   Df Sum Sq Mean Sq F value  Pr(>F)
June.Aug17.20$da_YR   1    284  283.62   8.813 0.00319 **
Residuals           359  11554   32.18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

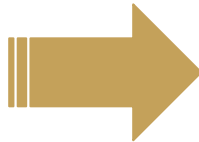# DOES NOT support H0

for any reasonable alphas

Methods

- 3 PAIRWISE HYPOTHESIS TESTINGS FOR TWO MEANS ON TWO CONSECUTIVE YEARS
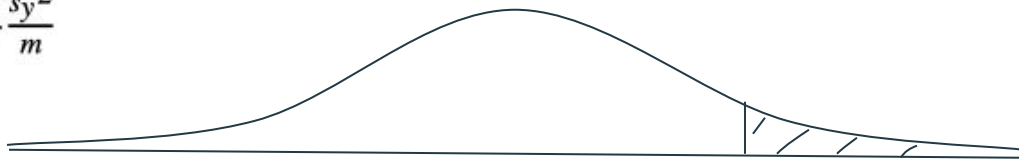
  - Sample size >=30
  - Variances unknown

  → Z distribution

X: later year dataset	Y: previous year dataset
n: sample size of X $\overline{X}$: sample mean of X s_x: sample standard deviation of X	m: sample size of Y $\overline{Y}$: sample mean of Y s_y: sample standard deviation of Y

Null hypothesis (H_0): $\mu_X = \mu_Y$

Alternative hypothesis (H_1): $\mu_X > \mu_Y$

Test statistics: $$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{s_x{}^2}{n} + \frac{s_y{}^2}{m}}}$$
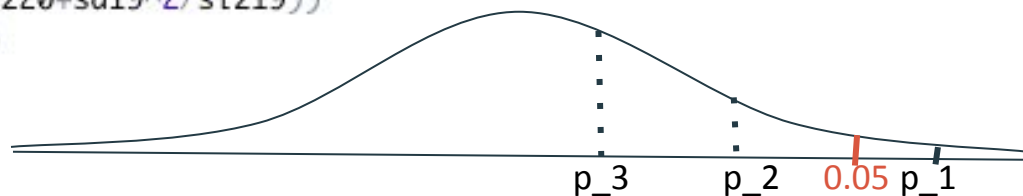
```
```

```r
#TEST1: X:18 ; Y:17
Z_1 = (mea18- mea17)/(sqrt(sd18^2/siz18+sd17^2/siz17))
p_1 = pnorm(Z_1, lower.tail = FALSE)


#TEST2: X:19 ; Y:18
Z_2 = (mea19- mea18)/(sqrt(sd19^2/siz19+sd18^2/siz18))
p_2 = pnorm(Z_2, lower.tail = FALSE)


#TEST2: X:20 ; Y:19
Z_3 = (mea20- mea19)/(sqrt(sd20^2/siz20+sd19^2/siz19))
p_3 = pnorm(Z_3, lower.tail = FALSE)
```

mea20	75.9919565217391
p_1	0.0266257007186893
p_2	0.441932791656761
p_3	0.173746592617741

Results

- ANOVA test: four means are not equal
- Hypothesis testing: only 2018 had a significant increase in temperature, but not for 2019 and 2020

Thank you so much for watching!

Data from: https://nassau-ny.weatherstem.com/hofstrasoccer

Help the polar bears at: https://www.wwf.org.uk/learn/wildlife/polar-bears

OR choose environmental friendly items when possible