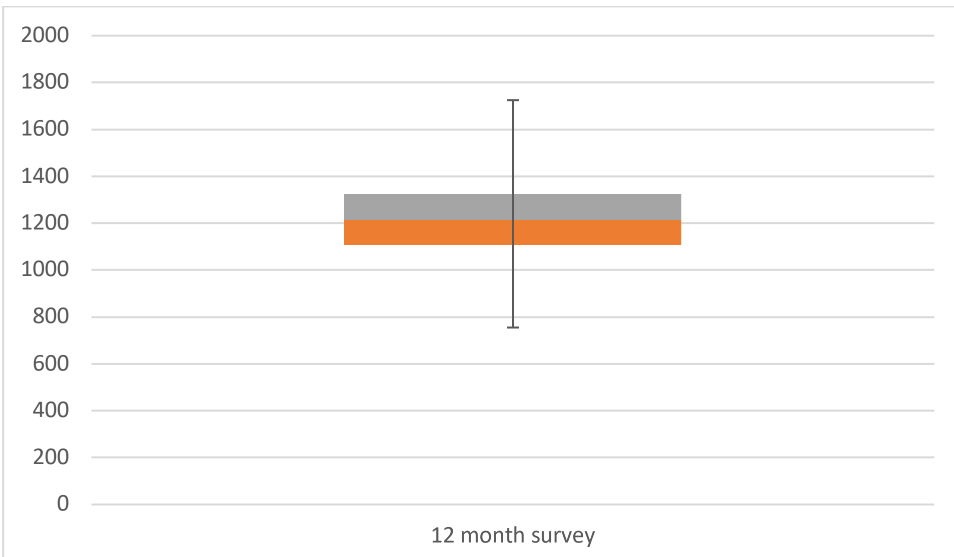
Maria Restrepo (mar981)

1. There are five potential outliers.

| Survey Rate | Z-Score |
|---|---|
| 755 | -2.91117 |
| 757 | -2.89843 |
| 1673 | 2.933516 |
| 1710 | 3.169086 |
| 1725 | 3.264588 |




2.

1. $H_o$: $\mu_1 = \mu_2$
   $H_A$: $\mu_1 \,!\!= \mu_2$
2. $\alpha$ =.05; $\alpha/2$ =.025
3. $\boldsymbol{t_{crit}} = \boldsymbol{t}_{\boldsymbol{df=n1+n2-2}}^{\boldsymbol{\alpha/2=.025}} = \boldsymbol{t}_{\boldsymbol{df=1053+535-2}}^{\boldsymbol{\alpha/2=.025}} = 1.959$
4. Reject $H_o$ if $|t_{obs}| > 1.959$

   $$t_{obs} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}};\ \text{where } s_{\bar{X}_1 - \bar{X}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$
5. 

   $$s_p = \sqrt{\frac{[(n_1 - 1)s_1^2] + [(n_2 - 1)s_2^2]}{(n_1 + n_2 - 2)}}$$

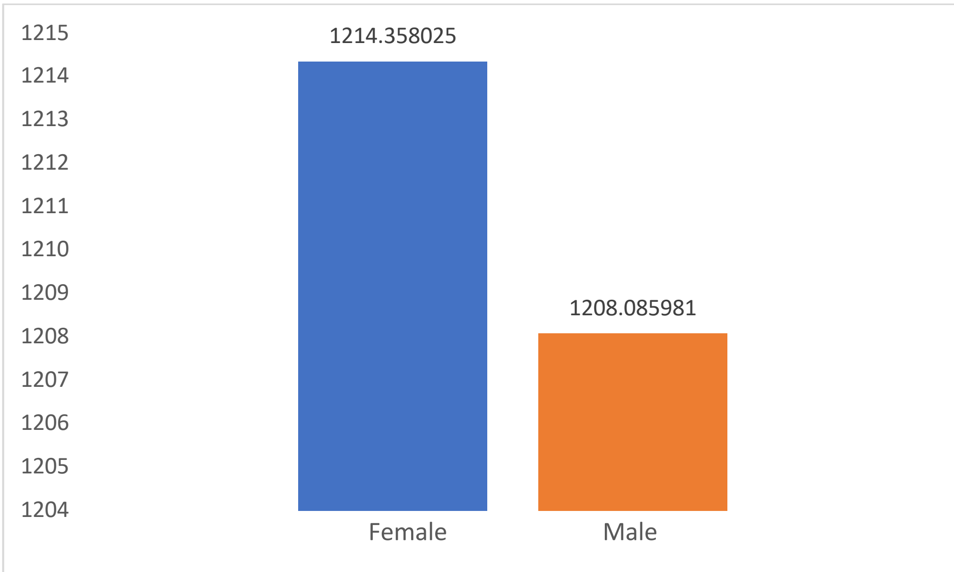
$$\boldsymbol{S_p} = \sqrt{\frac{\mathbf{(1052)(16031772)^2+(534)(150.815)^2}}{\mathbf{1053+535-2}}} == \sqrt{\mathbf{24676.42077}} = \mathbf{157.087303}$$

$$S_{Xbar1-xbar2} = 157.087303\sqrt{\frac{1}{1053}+\frac{1}{535}} = 8.340176776$$

$$t_{obs} = \frac{1214.358 - 1208.086}{8.340176776} = .7520224293$$

6. $t_{obs}(0.75) < t_{crit}$ (1.959) so fail to reject $H_o$. We conclude there is no statistical mean difference in survey score for male and female.

EXCEL OUTPUT:




3.

1. $H_o$: $\mu_1 = \mu_2$
   $H_A$: $\mu_1 \,!\!= \mu_2$
2. $\alpha$ =.05; $\alpha/2$ =.025
3. $t_{crit} = t_{df=n1+n2-2}^{\alpha/2=.025} = t_{df=1053+535-2}^{\alpha/2=.025} = 1.959$
4. Reject $H_o$ if $|t_{obs}| > 1.959$
5. $$t_{obs} = \frac{(\bar{X}_1 - \bar{X}_2)-(\mu_1-\mu_2)}{s_{\bar{X}_1-\bar{X}_2}};\ \text{where } s_{\bar{X}_1-\bar{X}_2} = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$
   $$s_p = \sqrt{\frac{[(n_1-1)s_1^2]+[(n_2-1)s_2^2]}{(n_1+n_2-2)}}$$

$$S_p = \sqrt{\frac{(1052)(.461635)^2+(534)(.49320286)^2}{1053+535-2}} == \sqrt{.22325563} = .472499$$

$$S_{Xbar1-xbar2} = .472499\sqrt{\frac{1}{1053}+\frac{1}{535}} = .025086211$$

$$t_{obs} = \frac{3.684786 - 3.57620561}{.025086211} = 4.328289$$

6. $t_{obs}(4.328289) > t_{crit}$ (1.959) so reject $H_o$. We conclude there is a statistical difference in survey score for male and female on manager's annual performance.

4.

1. $H_o$: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$
   $H_A$: $\mu_1$ != $\mu_2$ != $\mu_3$ != $\mu_4$ != $\mu_5$ != $\mu_6$
2. $\alpha$= 0.05
3. $F_{crit} = \boldsymbol{F^{\alpha=.05}_{dfn,dfd}} = \boldsymbol{F^{\alpha=.05}_{dfTR,dfE}} = \boldsymbol{F^{\alpha=.05}_{5,1582}} = 2.219$
   dfTR = k- 1 = 6-1 = 5
   dfErr = n-k = 1588-6 = 1582
4. Reject $H_o$ if $F_{obs} > F_{crit}$
5. $F_{obs}$ = 10.21 (from table)
6. $F_{obs}$ (10.21) > $F_{crit}$ (2.219) so reject Ho. Conclude that there is a statistically significant difference in the mean scores of the recruits' college major and the manager's annual performance review.

EXCEL OUTPUT:

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 11.2080113 | 5 | 2.24160225 | 10.2179437 | 1.1864E-09 | 2.21975448 |
| Within Groups | 347.057574 | 1582 | 0.219379 | | | |
| | | | | | | |
| Total | 358.265585 | 1587 | | | | |

5.

1. $H_o$: $\mu_D = 0$
   $H_A$: $\mu_D$ != 0
2. $\alpha$= 0.05: $\alpha$/2= 0.025
3. $t_{crit} = t_{df = 1589-1 = 1588}$ (t ^ $\alpha$ = .025) = 1.961
4. Reject $H_o$ if $|t_{obs}|$ (17.13622) > $t_{crit}$
5. $t_{obs}$ = -17.13622
6. $t_{obs}$(17.13622) > $t_{crit}$(1.961) so reject $H_o$. We conclude there is statistical change in the survey results between three months after they are initially hired and twelve months.

EXCEL OUTPUT:

t-Test: Paired Two Sample for Means

| | *three.month* | *twelve.month* |
|---|---|---|
| Mean | 1127.09509 | 1212.24496 |
| Variance | 20407.1939 | 24669.6646 |
| Observations | 1588 | 1588 |
| Pearson Correlation | 0.13075503 | |
| Hypothesized Mean Difference | 0 | |
| df | 1587 | |
| t Stat | -17.13622 | |

| | |
|---|---|
| P(T<=t) one-tail | 7.8837E-61 |
| t Critical one-tail | 1.64581435 |
| P(T<=t) two-tail | 1.5767E-60 |
| t Critical two-tail | 1.96145992 |

6.

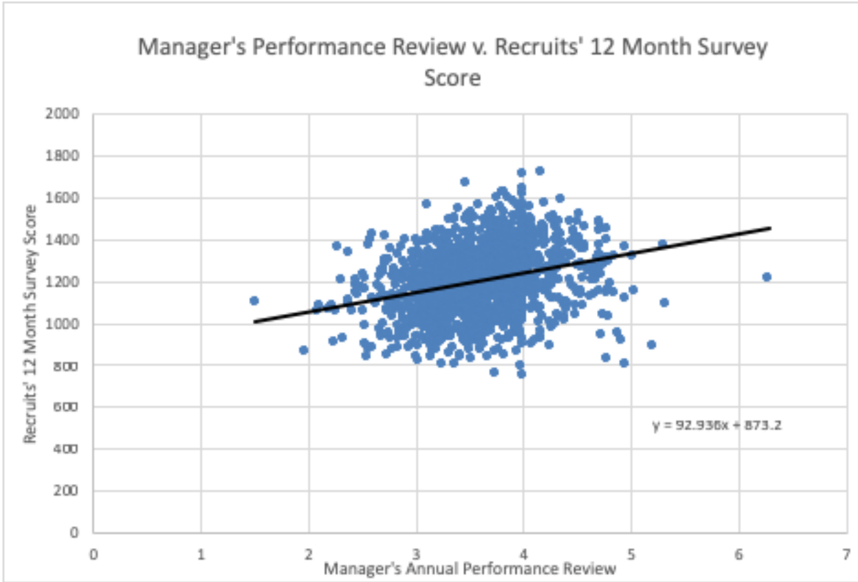
1. $H_o$: Manager's annual performance review does NOT predict the recruit's 12 month survey score.
   $H_A$: Manager's annual performance review does predict the recruit's 12 month survey score.
   OR
   $H_o$: $\beta_1 = 0$
   $H_A$: $\beta_1 \,!= 0$
2. $\alpha = .05$
3. $\boldsymbol{F_{crit} = F_{df=1,1586}^{\alpha=.05}}$ = 3.86 (from F table)
4. Reject $H_o$ if $F_{obs} > F_{crit}$ (3.86)
5. $F_{obs}$ = 136.11 (from ANOVA table)
6. $F_{obs}$ (136.11) > $F_{crit}$ (3.86), so reject $H_o$. Conclude that, manager's annual performance review does predict the recruit's 12-month survey score.

EXCEL OUTPUT:

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 1 | 3094352.99 | 3094352.99 | 136.110183 | 3.1173E-30 |
| Residual | 1586 | 36056404.7 | 22734.177 | | |
| Total | 1587 | 39150757.7 | | | |

Regression Equation: y = 0.0009x + 2.6173
Correlation: .28113495




7.

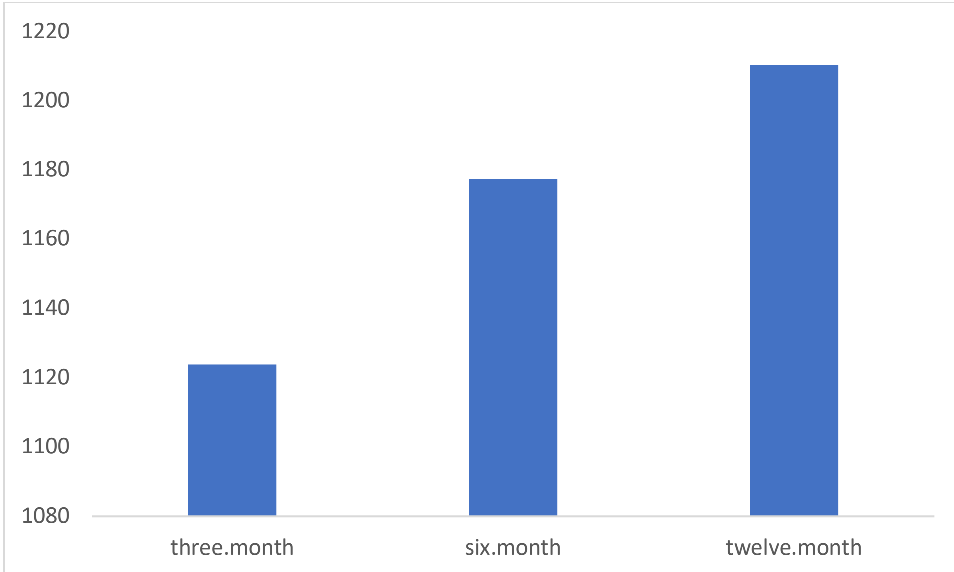
1. $H_o$: $\mu_1 = \mu_2 = \mu_3$
   $H_A$: $\mu_1 \mathrel{!=} \mu_2 \mathrel{!=} \mu_3$
2. $\alpha = 0.05$
3. $F_{crit} = = \boldsymbol{F^{\alpha=.05}_{2,1552}} = 3.00$
4. Reject $H_o$ if $F_{obs} > F_{crit}$
5. $F_{obs} = 74.86$ (from ANOVA table)
6. $F_{obs}$ (74.86) > $F_{crit}$ (3.00) so reject Ho. Conclude that there is a statistically significant difference on survey scores for three, six and twelve months.

EXCEL OUTPUT:

| | | | | |
|---|---|---|---|---|
| three.month | 777 | 873205 | 1123.81596 | 20347.7328 |
| six.month | 777 | 914864 | 1177.43115 | 22351.1116 |
| twelve.month | 777 | 940447 | 1210.3565 | 24187.2812 |

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Rows | 21172084.6 | 776 | 27283.6142 | 1.37787293 | 7.8552E-08 | 1.10662163 |
| Columns | 2965014.47 | 2 | 1482507.24 | 74.8693546 | 0 | 3.0015222 |
| Error | 30731548.9 | 1552 | 19801.2557 | | | |
| | | | | | | |
| Total | 54868648 | 2330 | | | | |




8.

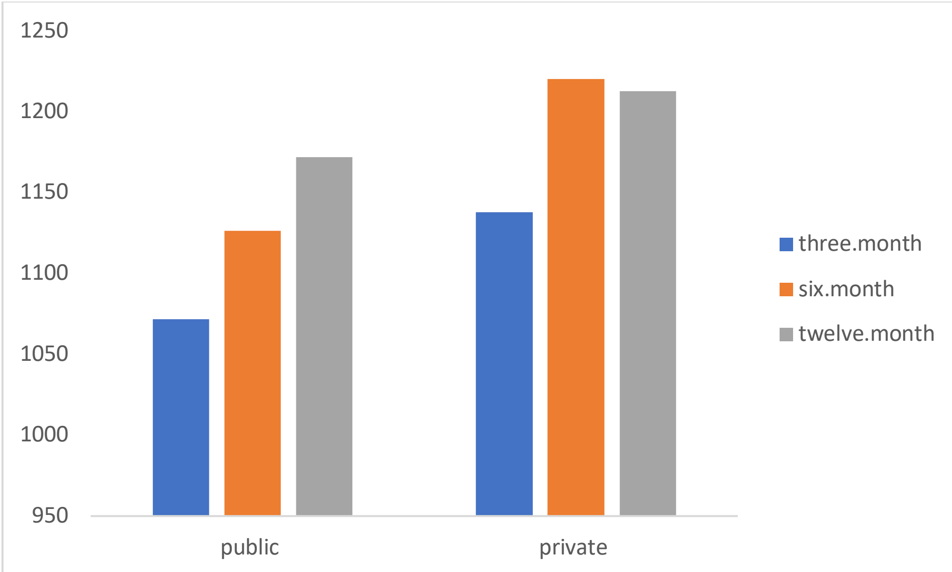
1st (the within) variable are the three times the survey was administered (the over time variable) and 2nd (the in btwn variable) is public or private institution.

1. $H_o$: There is NO significant interaction between survey score and institution.
   $H_a$: There is a significant interaction between survey score and institution.
2. $\alpha$ =.05
3. $F_{crit} = \boldsymbol{F^{\alpha=.05}_{dfn,dfd} = F^{\alpha=.05}_{dfTR,dfE} = F^{\alpha=.05}_{2,1734}}$= 3.00 (from ANOVA Table, Interactions row)
4. Reject $H_o$ if $F_{obs} > F_{crit}$ (3.00)
5. $F_{obs}$ = 4.80 (from ANOVA Table (F), Interactions row)
6. $F_{obs}$ (4.80) > $F_{crit}$ (3.00), so reject $H_{o.}$ Conclude that the interaction between survey score and institution is significant. This means there is a difference in survey scores for those who went to public and private institutions.

EXCEL OUTPUT:

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Sample | 1940874.01 | 1 | 1940874.01 | 91.2586322 | 4.1262E-21 | 3.8468272 |
| Columns | 2456434.34 | 2 | 1228217.17 | 57.7499717 | 0 | 3.00091379 |
| Interaction | 204319.71 | 2 | 102159.855 | 4.80348988 | 0.00831053 | 3.00091379 |
| Within | 36878434.9 | 1734 | 21267.8402 | | | |
| | | | | | | |
| Total | 41480062.9 | 1739 | | | | |




9. A

1. $H_o$: $\mu_1 = \mu_2$
   $H_A$: $\mu_1 \; != \mu_2$
2. $\alpha = 0.05$
3. $F_{crit} = = \boldsymbol{F^{\alpha=.05}_{1,1586}} = 3.84$
4. Reject $H_o$ if $F_{obs} > F_{crit}$
5. $F_{obs} = 3.15$ (from ANOVA table)
6. $F_{obs}$ (3.15) < $F_{crit}$ (3.84) so fail reject Ho. Conclude that there is no statistically significant difference on 12-month survey scores for the recruits that attended HBCU institutions and those that didn't.

EXCEL OUTPUT:

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 77534.636 | 1 | 77534.636 | 3.14716635 | 0.07625076 | 3.84732872 |
| Within Groups | 39073223.1 | 1586 | 24636.3323 | | | |
| | | | | | | |
| Total | 39150757.7 | 1587 | | | | |

9.B

1. $H_o$: $\mu_1 = \mu_2$
   $H_A$: $\mu_1 \; != \mu_2$
2. $\alpha = 0.05$
3. $F_{crit} = = \boldsymbol{F^{\alpha=.05}_{1,1586}} = 3.84$
4. Reject $H_o$ if $F_{obs} > F_{crit}$
5. $F_{obs} = .9788$ (from ANOVA table)

6. $F_{obs}$ (.9788) < $F_{crit}$ (3.84) so fail reject Ho. Conclude that there is no statistically significant difference on 12-month survey scores for the recruits that attended H.S.I institutions and those that didn't.

EXCEL OUTPUT:

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 24147.0686 | 1 | 24147.0686 | 0.97880318 | 0.32264537 | 3.84732872 |
| Within Groups | 39126610.6 | 1586 | 24669.9941 | | | |
| | | | | | | |
| Total | 39150757.7 | 1587 | | | | |

## 9.C

1. $H_o$: HBCU institutions and region are independent.
   $H_A$: HBCU institutions and region are NOT independent.
2. α= 0.05
3. $\boldsymbol{X^2_{crit}} = \boldsymbol{X^{2\ \alpha=.05}_{df=(2-1)(4-1)=3}}$= 7.815 (from table)
4. Reject $H_o$ if $X^2_{obs}$> $X^2_{crit}$ (7.815)
5. $X^2_{obs}$ = 141.16 (from Excel)
6. $X^2_{obs}$ (141.16) > $X^2_{crit}$ (7.815), so reject $H_o$. Conclude that HBCU institutions and region are NOT independent.

EXCEL OUTPUT:

| *Expected* | Northeast | South | Midwest | West | |
|---|---|---|---|---|---|
| **Is HBCU** | 5.3 | 54.8 | 50.8 | 50.2 | **167** |
| **Is not HSBU** | 44.7 | 466.2 | 483.2 | 427 | **1421** |
| | **50** | **521** | **540** | **477** | **1588** |

| **Count of inst.region** | **Column Labels** | | | | |
|---|---|---|---|---|---|
| **Row Labels** | **1** | **2** | **3** | **4** | **Grand Total** |
| 0 | 50 | 354 | 540 | 477 | 1421 |
| 1 | | 167 | | | 167 |
| **Grand Total** | **50** | **521** | **540** | **477** | **1588** |

## 10.

Are the recruits satisfied?

To element some noise variables, we were able to determine that gender does not significantly affect the recruit's survey scores. We were also able to determine there is a significant change in survey response after the three months after hire survey. It was also concluded, based on the recruiters' major, the manager rating would differ.

Are their manager's reviews of their performance in sync with their own survey responses?

For the recruiter's own survey of success gender did not play a role, but when it came to the survey scores for the manager, the gender of the recruiters was statistically significant. We were also able to determine that the 12-month recruiter survey score relies on the manager's yearly performance review survey. This implies a complimentary relationship between the manager and the recruits.

Also, where should we go to recruit if we want to increase diversity in our workforce?

If we want to increase diversity in the workforce, then it is necessary to target specific areas. Based on the hypothesis testing done in Question 9C, we were able to determine that region plays a big role in receiving recruits from a certain background. Historically Black Colleges are located in a very specific region, so in order to raise diversity there has to be a targeted effort in the South.

## 11.

An additional analysis we can consider is testing to see if recruits' survey result gets significantly better or worse after they are initially hired.

We can also run tests to see:

- If there is a significant difference in SAT scores depending on the level of competitiveness the schools have?
- Does the primary language spoken at home predict SAT scores?
- Is the level of competitiveness associated with or independent from region?
- Is the primary language spoken at home associated with or independent form the region?

I believe conducting these tests would a provide a bigger picture of the recruits we are hiring and obstacles they might face.