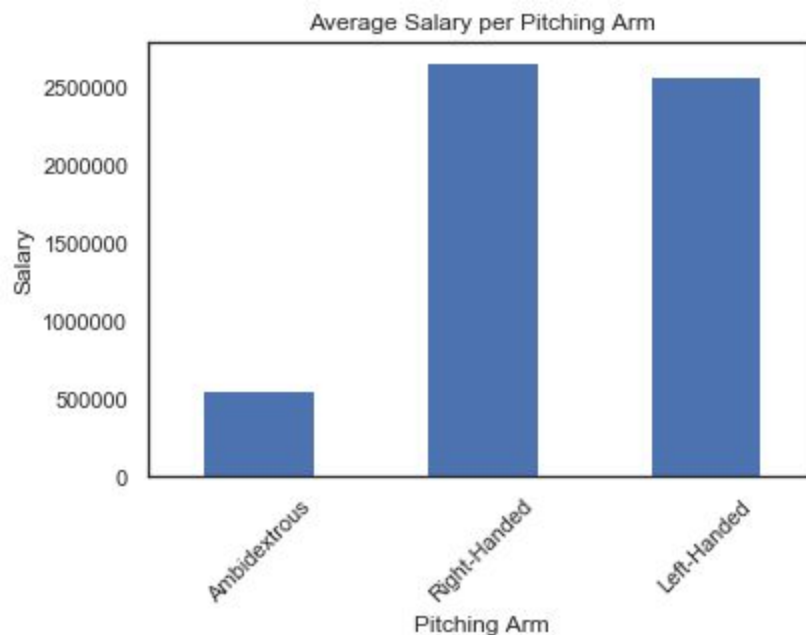


Statistical Inference



Is there a difference between the salaries paid towards right handed and left handed pitchers?

H_0 : The salaries for both left and right handed pitchers are the same.

H_1 : The salaries for left and right handed pitchers are different.

An independent-samples t-test was conducted to compare the salaries paid to left and right handed pitchers, with an alpha level set at .05. Results of the independent sample t-tests indicated that there were not significant differences in salaries paid to pitchers who threw with their left or right arm, ($t(7689) = -0.75$, $p = 0.77$). Specifically, our results suggest that pitchers are paid the same regardless of throwing arm.

```
# manual t-test
lefties = df_pitcher.adj_salary_filled[df_pitcher['Throws'] == 'L']
righties = df_pitcher.adj_salary_filled[df_pitcher['Throws'] == 'R']

alpha= 0.05
N = len(lefties) + len(righties)

#Calculate mean
```

```

lefties_mean = np.mean(lefties)
righties_mean = np.mean(righties)

#Calculate variance
lefties_var = np.var(lefties)
righties_var = np.var(righties)

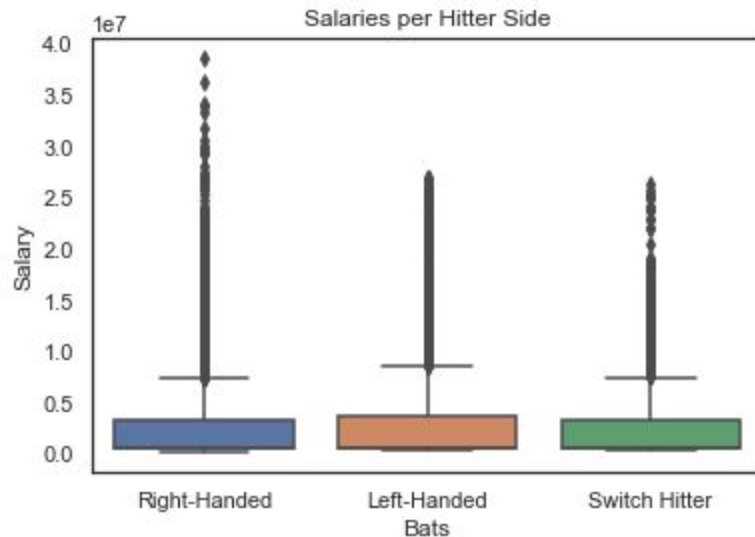
#calculate std
# std = np.sqrt((lefties_var + righties_var)/2)
std_lefties = np.std(lefties, ddof=1)
std_righties = np.std(righties, ddof=1)
len_lefties, len_righties = len(lefties), len(righties)
std = np.sqrt((((len_righties - 1)*(std_righties**2)) + ((len_lefties - 1) *
(std_lefties**2)))) / (N - 2))

#Calculate the t-statistics
t = (righties_mean - lefties_mean)/(std*np.sqrt((1/len_righties +
1/len_lefties)))
df = N - 2

#calculate p value
p_value = 1 - stats.t.cdf(t,df=df)
print(p_value, t)
0.7743908173193755 -0.7534237812415734

# compare samples with t-test from SciPy
ttest,pval = ttest_ind(righties,lefties)
print("p-value",pval, "t-test",ttest)
if pval <0.05:
    print("We reject the null hypothesis.")
else:
    print("We fail to reject the null hypothesis.")
p-value 0.45121836536138527 t-test -0.7534237812413467
We fail to reject the null hypothesis.

```



Is there a difference between the salaries paid towards right handed and left handed batters?

H_0 : The salaries paid to both left and right handed hitters the same.

H_1 :The salaries paid to left and right handed hitters are different.

An independent-samples t-test was conducted to compare the salaries paid to left and right handed batters, with an alpha level set at .05. Results of the independent sample t-tests indicated that there were not significant differences in salaries paid to hitters who batted with their left or right side, ($t(5973) = 1.35$, $p = 0.17$). Our results suggest that hitters are paid the same regardless of batting side.

```
lefties = df_hitter.adj_salary_filled[df_hitter['Bats'] == 'L']
righties = df_hitter.adj_salary_filled[df_hitter['Bats'] == 'R']
```

```
ttest,pval = ttest_ind(lefties,righties)
print("p-value",pval, 'ttest', ttest)
if pval <0.05:
    print("We reject the null hypothesis.")
else:
    print("We fail to reject the null hypothesis.")
```

p-value 0.17474496847305468 ttest 1.3572755073329725
We fail to reject the null hypothesis.