Final Project Report

Introduction:

In just over a century, air travel has grown from nothing into a booming industry, essential to global commerce, and a significant part of our everyday lives. As the worldwide population and economy continues to grow, so does the demand for reliable air transportation. One of the world's largest hubs is located nearby at as Hartsfield-Jackson Atlanta International Airport. As graduate students in the Aerospace Systems Design Lab at Georgia Tech, the authors have experience and interest in studying aviation and airline operations. We are all frequent flyers through the Atlanta airport, and the performance of the airlines and airspace system often affects our travels. Therefore, for this project, we have chosen to investigate several questions regarding flight delay times and airline operational consistency.

ATL is one of the busiest airports in the world, and serves as a hub for two major airlines, Delta and AirTran. The data chosen for this study comes from the Bureau of Transportation Statistics (BTS) Research and Innovative Technology Administration (RITA)[1]. The BTS acts within the U.S. Department of Transportation (DOT) to provide data collection and analyses for the country's many modes of travel. This source provides departure and arrival time information on flights within the U.S. since 1987. However, the amount of available data is tremendous, spanning many years and airports. For the purposes of this project, the scope of our study will be limited to flights leaving ATL in 2012.

The specific information available within the dataset includes, but is not limited to: flight number, carrier, scheduled and actual departure and arrival time, delay time, origin and destination city, and reasons for delays. The source website provides this data for specific years, months, or geographic region, and allows for easy selection of the most relevant information to be downloaded. For this project, we have compiled data on all 392,015 flights that departed from ATL in 2012, and this is taken as our population of interest.

The amount of raw data available from this source is enormous, and could be useful in studying many different aspects of air travel delays. However, we are most interested in a few specific questions regarding air-travel through Atlanta during the year 2012. Although the entire population is available, we have approached these studies generally assuming unknown population statistics, and instead relied on inference from various sample statistics and analysis. The following are major questions of interest:

1. Are the average departure and arrival delays for the entire airport significant?
2. What are the most common causes of departure delays?
3. What time of day are delays more likely to occur?
4. Which airlines are most consistently on-time?
5. What time of the year are departure delays most likely to occur?
6. Does the number of cancelled flights in a given day follow a Poisson distribution?
7. Which carrier has fewer cancelled flights – Delta or AirTran?

By using methods learned in this course, we hope to answer these questions and discover others that could be investigated in the future. Analysis of the data at hand could lead to some enlightening results about air travel from Atlanta and potentially influence our future travel planning decisions, from choosing the best airline to deciding on the best time to fly to avoid unforeseen delays.

General Assumptions:

For all analyses throughout this report, the following assumptions are made, unless otherwise noted:

- Population variance is unknown
- When comparing multiple samples, their population variances are considered unequal
- $\alpha = 0.05$

Question I: Are the average departure and arrival delays for the entire airport significant?

A. Test for Average DEPARTURE Delay

The first investigation into flight delays at Atlanta International Airport consists of a simple hypothesis test on the mean departure delay time. From a random sample of 10,000 flights across the entire spectrum of carriers, dates, and times throughout the year 2012, 43 flights were cancelled, leaving 9957 to use in the test. Although data for the entire population of interest is available, for this test we make the more realistic assumption of unknown population variance. We also assume that any flight departure "delay" time that is reported as negative should be treated as zero, since our study focuses on flight tardiness. These assumptions lead us to adopt a one-sided hypothesis test using the *t*-statistic.

μ : Population average departure delay, in minutes, from ATL in 2012
X_{bar} : Sample average departure delay, in minutes, from ATL in 2012
Assume: Negative "delays" ➔ 0

H_0: $\mu = 0$; On average, flights from Atlanta in 2012 were not delayed at all.
H_a: $\mu > 0$; On average, there was some level of departure delay.

Sample Statistics:

$n = 9957$, $X_{bar} = 8.982$, $S = 25.17$

Test Statistic	95% C.I.	P-value
$t_0 = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} = \frac{8.982-0}{25.17/\sqrt{9957}} = 35.61$ $t_{0.05,9956} \approx z_{0.05} = 1.645$ $t_0 > t_{0.05,9956}$	$\left[\bar{X} - z_{0.05}\frac{S}{\sqrt{n}}, \infty\right]$ $[8.567, +\infty)$	$[1 - T_{9956}(\lvert t_0 \rvert)]$ ≈ 0.00

**A *t*-distribution with more than 30 degrees of freedom approximates a normal distribution, so we can use the *z*-statistic.

⇨ **Reject H0**

The test statistic is far greater than the critical value, the hypothesized mean falls outside of our 95% confidence interval, and the p-value is almost zero, thus we must reject the null hypothesis. There was clearly a non-zero average departure delay time from ATL in 2012. The large sample used here suggests a mean departure delay closer to 10 minutes.

Power of Test:

To investigate the power of this test, consider that the true population mean can be determined based on the complete set of data collected for flights from ATL in 2012. We will still assume that the population variance is unknown, but given a true mean delay time greater than zero, what is the probability of failing to reject H_0 when H_0 is false?

Population Statistics:

N = 388,958, $\mu_1 = 9.140$, $\delta = \mu_{1} - \mu_0 = 9.140$

The previous hypothesis test with n = 9957 has Type II error:

$$\beta = T_{9956}\left(t_{0.05,9956} - \frac{\delta\sqrt{n}}{S}\right) \approx \Phi\left(1.645 - \frac{9.14\sqrt{9957}}{25.17}\right) = \Phi(-34.59) \approx 0$$

Such a large sample size makes the test very powerful, and in this case the data is readily available and easy to sample, so obtaining a large sample size is not costly. In general, what is the minimum sample size needed to design a test for this data with Type II error less than 0.10?

$$n = \frac{\left(t_{0.05,9956} + t_{0.10,9956}\right)^2 S^2}{\delta^2} = \frac{(1.645 + 1.282)^2\, 25.17^2}{9.14^2} = 64.97 \approx 65$$

A random sample of only 65 flights will yield a hypothesis test for mean departure delay with Power = 0.90.

B. Test for Average ARRIVAL Delay

While we have shown that the average departure delay in 2012 was nonzero, that does not necessarily imply that the average time of arrival is also late. The same random sample of 10,000 flights used in the previous test is considered again, but this time 28 additional flights were diverted and never reached their intended destination. The remaining useful sample of 9929 flights were used to perform a two-sided hypothesis test on mean arrival delay time, again assuming unknown population variance and using the *t*-statistic.

μ : Population average arrival delay, in minutes, from ATL in 2012
X_{bar} : Sample average arrival delay, in minutes, from ATL in 2012
Assume: Consider both negative and positive "delays."

H_0: $\mu = 0$; On average, flights arrived at their destination right on time.
H_a: $\mu \neq 0$; On average, flights arrived either early or later than scheduled.

Sample Statistics:

n = 9929, $X_{bar} = 0.119$, S = 28.511

Test Statistic	95% C.I.	P-value
$t_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{0.119 - 0}{28.511/\sqrt{9929}} = 0.416$ $t_{0.025,9929} \approx z_{0.025} = 1.96$ $t_0 < t_{0.025,9929}$	$\left[\bar{X} - z_{0.025}\frac{S}{\sqrt{n}}, \bar{X} + z_{0.025}\frac{S}{\sqrt{n}}\right]$ [-0.442, 0.680]	$2\left[1 - T_{9956}(\lvert t_0 \rvert)\right]$ = 0.667

⇨ **Fail to Reject H0**

The results indicate that the test statistic is less than the critical value, the hypothesized mean falls within the 95% confidence interval, and the p-value is greater than α, so the null hypothesis is NOT rejected. We cannot say that average arrival delay is nonzero. This suggests that on average, even when departing late, flights still tend to arrive on time.

<u>Question II:</u> What are the most common causes of arrival delays?

Since June of 2003, airlines were required to report flight data to the Bureau of Transportation Statistics including causes of delays and cancellations[2]. The causes of delays are classified by five categories, defined by the Air-Carrier On-Time Reporting Advisory Committee. The causes of delays include the following: air carrier delay, weather delay, National Aviation System (NAS) delay, security delay and late aircraft delay. Air carrier delay is defined by any delay caused by circumstances controlled by the airline. It includes delays caused by maintenance, fueling, aircraft cleaning, loading luggage, along with many others. Weather delays are delays caused by actual or forecasted significant meteorological conditions including hurricanes and blizzards. NAS delays are delays attributed to the National Aviation system which correspond to airport operations and air traffic control. It also includes non-extreme weather delays. Security delays are due to any evacuation caused by security violations and delays in security screenings including long screening lines which are greater than 29 minutes. Late aircraft delays are caused from an aircraft arriving late from the previous flight and departing late for the current flight.

In analyzing the most common causes for arrival delays, a one-way Analysis of Variance (ANOVA) was implemented. Five random samples of 1000 flight data points were selected from the population data set, each to collect data for a different cause of arrival delay. ANOVA tests the following:

In order to perform ANOVA, two assumptions need to hold: the residuals must be normally and independently distributed (NID) and homogeneity of the variances across the samples (treatments have equal variances). The assumptions do not hold in this case. A normality test for the residuals resulted in a p-value of less than 0.005 rejecting the null hypothesis stating that the residuals are normal. The histogram of the residuals given by

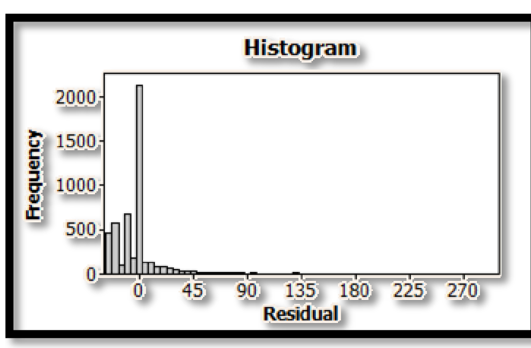


Figure 1. Histogram of residuals

Figure 1 and the normal probability plot given by Figure 2 illustrate the normality assumption does not hold. In testing the homogeneity of variance between the treatments, two tests were conducted including the Bartlett's test and the Levene's test. Both tests resulted in a p-value of 0.00 rejecting the null hypothesis stating that the treatments having equal variances. The results of the tests are illustrated by Bonferroni confidence intervals for standard deviations given by Figure 3. Since the assumptions required for ANOVA do not hold, the Kruskal-Wallis method was used to test the differences between the samples.

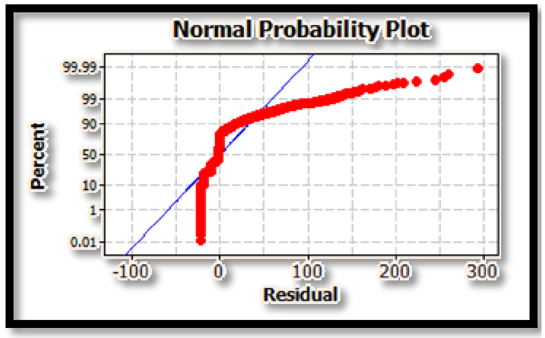


Figure 2. Normal probability plot for residuals

The Kruskal-Wallis is a non-parametric method, analogous to the one-way ANOVA problem. Instead of testing for equal means, the method tests for equal medians. The null and alternative hypotheses for this case are given below:

The test statistic, K_0, is a rank based statistic which follows a chi-squared distribution, χ^2. It is calculated with in two different manners depending on if there are repeated observations within the test data. If there are no repeated observations, the problem has "no ties". In this case the test statistic is calculated by the following:

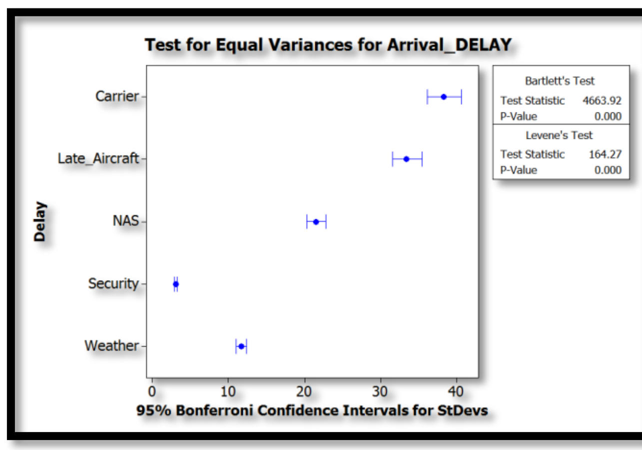


Figure 3. Test for equal variances for causes of arrival delays

where, C is the number of samples (treatments), n_i is the number of observations in the i^{th} sample, N is the total number of observations in all samples combined, and R_i is the sum of the ranks in the i^{th} sample. If the problem includes repeated observations, the problem has "ties". In this scenario, K_0 is calculated by the following equation:

where G is the number of grouping of the tied ranks and t_i is the number of tied observations within a group. If the test statistic, K_0 is larger than $\chi^2_{\alpha,C-1}$, then the null hypothesis is rejected. Besides using the test statistic to test the hypothesis, the p-value can be utilized. The p-value is calculated with the following probability:

If the p-value is less than the specified significance level, α, then the null hypothesis is rejected. In this study, the calculated test statistic is $K_0=1243.18$ which is greater than $\chi^2_{0.05,4}=9.488$ so the null hypothesis is rejected. Similarly, the p-value was calculated to be zero, which is smaller than the significance level, α =0.05. Once again the null hypothesis is rejected and the medians of the different sources of arrival delays contain some difference.

Two-sided confidence intervals were calculated for the mean of each sample to find out which causes are the main sources for arrival delays. The population variance is unknown so the t-distribution was used in constructing the confidence intervals. The following 2-sided 95% confidence was used to calculate the confidence intervals:

The confidence intervals and the descriptive statistics used in calculating the confidence intervals are given in Appendix A. The mean, standard deviation, and confidence intervals are all measured in minutes. The confidence intervals are illustrated in Figure 4 with the marker specifying the value of the sample mean.

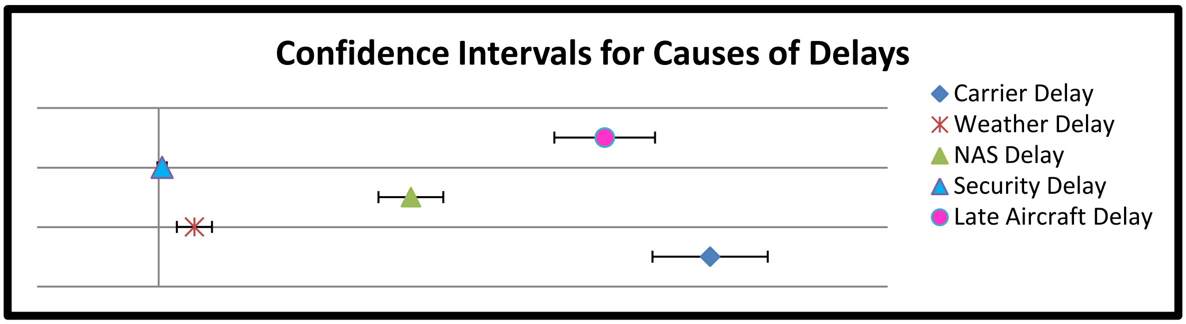


Figure 4. 95 % Confidence intervals for causes of arrival delays. Markers indicate the sample means.

In comparing all of the confidence intervals for the causes of arrival delays, it can be seen that carrier delays are mainly responsible for the arrival delays. Late aircraft delay is another main cause for arrival delays. On the other hand, security delays have the lowest mean value and a small confidence interval indicating low variability in the sample. This makes sense as security delays are quite rare and do not account for any delays in most flights.

<u>Question III:</u> What time of day are arrival delays more likely to occur?

In investigating this problem, the times of day were first defined. One complete day (24 hour period) was divided into four 6-hour periods. The times of day are defined in Table 1. Late-night flights include the red-eye flights which carry over into the next morning. Four random samples were collected from the population data set. Each sample represents a time of day with 1000 data points.

Time of Day	Morning	Midday	Evening	Late-night
Time	5:00am-11:00am	11:00am-5:00pm (11:00-17:00)	5:00pm-11:00pm (17:00-23:00)	11:00pm-5:00am (23:00-5:00)

Table 1. Time of day time definitions

A one-way Analysis of Variance (ANOVA) was implemented testing the following null and alternative hypotheses:

Two assumptions were tested to validate the results of ANOVA including a test for the normality of the residuals and tests for equal variance across all treatments or in this case, time of day. The normality test for the residuals resulted in a p-value of less than 0.005 rejecting the null hypothesis stating that the residuals are normal. The histogram and the normal probability plot of the residuals illustrate the normality assumption does not hold. The plots are shown in Figure 7 and Figure 6, respectively. The Bartlett's test and the Levene's test were used in testing if the there are equal variance across all the times of day. Both tests resulted in a p-value of 0.00 rejecting the null hypothesis that there are equal variances across all times of day. The results of the tests are illustrated by Bonferroni confidence intervals for standard deviations given by Figure 7. Since the assumptions required for ANOVA do not hold, the Kruskal-Wallis method was used to test the differences between the samples with the null hypothesis stating that the median of the arrival delays for each time of day are all equal. The calculated test statistic is $K_0=193.38$ which is greater than $\chi^2_{0.05,3}=7.815$ so the null hypothesis is rejected. Similarly, the p-value was calculated to be zero, which is smaller than the significance level, α =0.05. Once again the null hypothesis is rejected and the median of the arrival delays for each time of delay are not all equal.

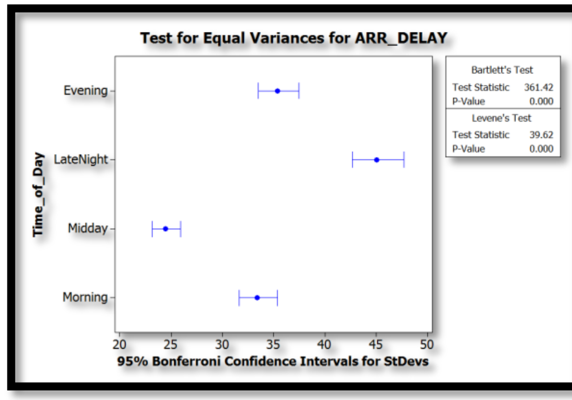


Figure 6. Test for equal variances for times of day

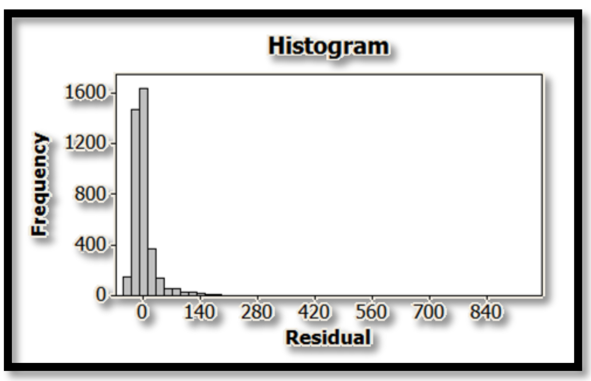


Figure 5. Histogram of residuals

Two-sided confidence intervals were calculated for the mean of each sample to find out which times of day are related to the highest and lowest arrival delays. The population variance is unknown so the t-distribution was used in constructing the confidence intervals. The

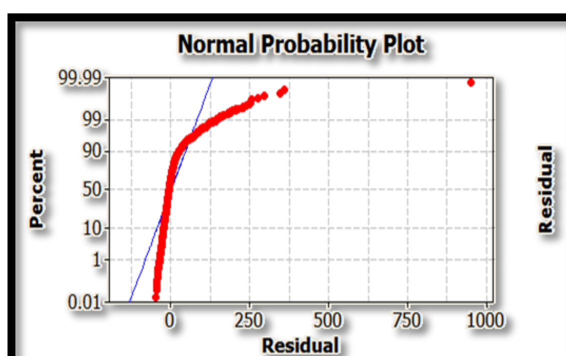


Figure 7. Normal probability plot for residuals

following 2-sided 95% confidence was used to calculate the confidence intervals:

The results are shown in Figure 8. See Appendix B for the values used in the calculations and the actual values of the confidence factors.

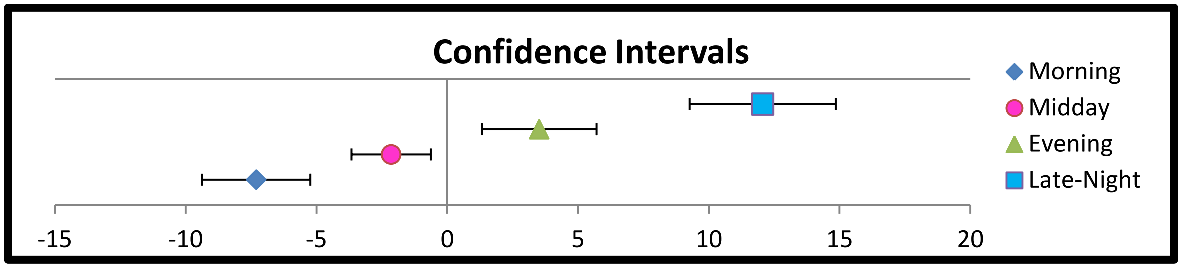


Figure 8. 95% Confidence intervals for the times of day arrival delays. Markers specify the sample means.

The morning and afternoon flights tend to arrive early while the evening and late-night flights tend to arrive late. The average arrival delays increase with the later times of day, ranging from an average early arrival of approximately 10 minutes to an average late arrival of approximately 15 minutes. The confidence intervals increase with the later times of day with the exception of decreasing going from morning flights to afternoon flights. This can be explained by the scatter plot of the arrival delays given the time of day illustrated in Figure 9. The arrival delay times increase over the course of a day, as delays can accumulate over different flights. This accumulation of delay can be seen to carry over from the red-eye flights into the early morning. This explains the larger variation in the morning compared to the afternoon.

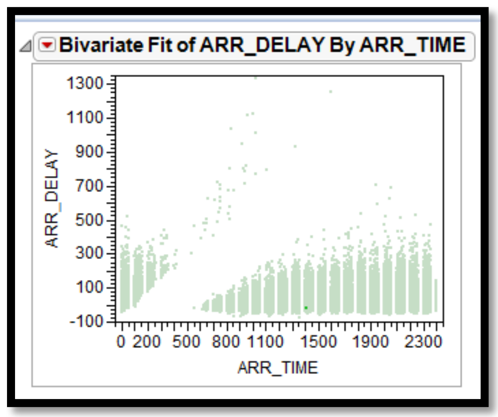


Figure 9. Scatter plot of arrival time vs. arrival Delay

<u>Question IV</u>: Are average delays different between airlines?

A. Test for Difference in DEPARTURE Delay between Carriers

To investigate differences in specific airline operations and delay times, a Kruskal-Wallis one-way analysis of variance is performed. Because the population distribution of flight delay times from ATL in 2012 is not Normal, the standard ANOVA approach is not appropriate, and this alternative method tests for differences in sample median. A new random sample is selected of 300 flights from each of the 12 airlines in our population dataset (3600 total). We again assume unknown and unequal population variances, and that any negative departure delay should be treated as zero. A summary of these test results from Minitab are given below.

m_i : Population median departure delay, in minutes, from ATL in 2012 for given carrier
$\tilde{X}_i$: Sample median departure delay, in minutes, from ATL in 2012 for given carrier

Response: Departure Delay Time
Factor: Carrier
Assume: Independent, non-normal sample distributions between carriers.
Negative "delays" ➔ 0

H_0: $m_1 = m_2 = \ldots = m_{12}$
H_a: At least one is different

Kruskal-Wallis Test: DEP_DELAY_Positive versus CARRIER

```
H = 80.30  DF = 11  P = 0.000
H = 115.78  DF = 11  P = 0.000  (adjusted for ties)
```

Test Statistic and P-value:

The Kruskal-Wallis test statistic, H or K, may be compared to critical values from the Chi-Square distribution corresponding to 11 degrees of freedom, one less than the number of carriers.

$$K_0 = 80.30 \qquad \chi^2_{11,\,0.05} = 19.675$$
$$K_0 > \chi^2_{11,\,0.05}$$
$$\text{P-value} = \Pr(\chi^2_{11} \geq K_0) \approx 0$$

➔ **Reject H0**

The test statistic is far greater than the critical value and the p-value is very near zero, thus we must reject the null hypothesis. The 12 carriers sampled have different median departure delays. Next, we perform the same test on average arrival delays across the different carriers. The same assumptions hold, but this time early-arriving flights are left as negative numbers.

B. Test for Difference in ARRIVAL Delay between Carriers

Response: Arrival Delay Time
Factor: Carrier

H_0: $m_1 = m_2 = \ldots = m_{12}$
H_a: At least one is different

Kruskal-Wallis Test: ARR_DELAY versus CARRIER

```
H = 112.25  DF = 11  P = 0.000
H = 112.30  DF = 11  P = 0.000  (adjusted for ties)
```

Test Statistic and P-value:

$$K_0 = 112.25 \qquad \chi^2_{11,\,0.05} = 19.675$$
$$K_0 > \chi^2_{11,\,0.05}$$
$$\text{P-value} = \Pr(\chi^2_{11} \geq K_0) \approx 0$$

➔ **Reject H0**

Once again, we the test results indicate that we must reject the null hypothesis. Not all of the airlines sampled had the same median arrival delay in 2012. The limitations of the Kruskal-Wallis method prevent us from determining the confidence intervals for all airlines at once, but we can focus on two of the largest carriers in ATL to investigate more.

C. Test for Difference in DEPARTURE Delay between Delta and AirTran

In response to the results of the previous Kruskal-Wallis tests, we choose to perform an individual *t*-test on the difference in population mean delay times. Using the same data sample of 300 flights for each airline and the same assumptions as above, a two-sided hypothesis test using the *t*-statistic is conducted.

Δ : Difference in population average departure delays, in minutes, between Delta and AirTran
X_{bar1} - X_{bar2}: Difference in sample average departure delays between Delta and AirTran
Assume: Negative "delays" ➔ 0

H_0: $\Delta = 0$; On average, Delta and AirTran experience the same level of departure delay
H_a: $\Delta \neq 0$; On average, Delta and AirTran experience different departure delays

Sample Statistics:

Delta:
n =300, X_{bar} = 8.17, S = 23.13

AirTran:
n =300, X_{bar} = 3.37, S = 11.12

Test Statistic**	95% C.I.	P-value
$t_0 = \dfrac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} = 3.24$ $\nu = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}} = 322.9 \approx 323$ $t_{0.025,323} \approx z_{0.025} = 1.96$ $t_0 > t_{0.025,323}$	$\left[(\bar{X}_1 - \bar{X}_2) - z_{0.025} S_{\hat{\Delta}}, (\bar{X}_1 - \bar{X}_2) + z_{0.025} S_{\hat{\Delta}}\right]$ $S_{\hat{\Delta}} = \sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}$ [2.098, 7.502]	$2\left[1 - T_{323}(\lvert t_0 \rvert)\right]$ ≈ 0.0022

**For a *t*-test with unknown and unequal variances for difference in populations, we must calculate the degrees of freedom differently

➔ Reject H0

The test results indicate that AirTran, on average, experiences less departure delay than Delta. Performing the same procedure for arrival delays yields the following results:

D. Test for Difference in ARRIVAL Delay between Delta and AirTran

Δ : Difference in population average arrival delays, in minutes, between Delta and AirTran
X_{bar1} - X_{bar2}: Difference in sample average arrival delays, in minutes, between Delta and AirTran

H_0: $\Delta = 0$; On average, Delta and AirTran experience the same level of arrival delay
H_a: $\Delta \neq 0$; On average, Delta and AirTran experience different arrival delays

Sample Statistics:

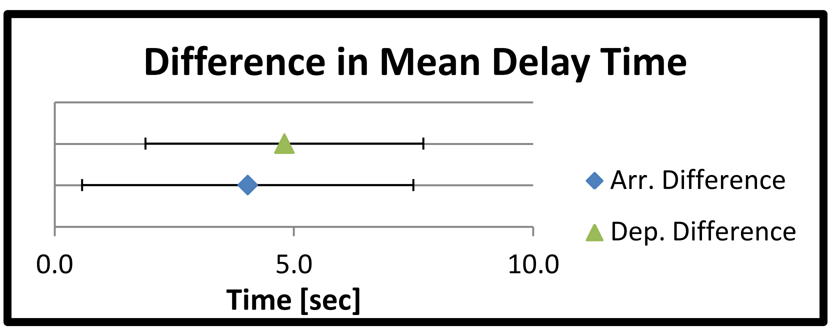
Delta:	AirTran:
n =300, X_{bar} = -0.607, S = 25.7	n =300, X_{bar} = -4.64, S = 16.6

Test Statistic**	95% C.I.	P-value
$t_0 = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{S_1^{\ 2}}{n_1} + \frac{S_2^{\ 2}}{n_2}}} = 2.28$ $\nu = \frac{\left(\frac{S_1^{\ 2}}{n_1} + \frac{S_2^{\ 2}}{n_2}\right)^2}{\frac{\left(\frac{S_1^{\ 2}}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^{\ 2}}{n_2}\right)^2}{n_2 - 1}} = 511.5 \approx 512$ $t_{0.025,512} \approx z_{0.025} = 1.96$ $t_0 > t_{0.025,512}$	$\left[(\bar{X}_1 - \bar{X}_2) - z_{0.025} S_{\hat{\Delta}}, (\bar{X}_1 - \bar{X}_2) + z_{0.025} S_{\hat{\Delta}}\right]$ $S_{\hat{\Delta}} = \sqrt{\frac{S_1^{\ 2}}{n_1} + \frac{S_2^{\ 2}}{n_2}}$ [0.571, 7.495]	$2\left[1 - T_{323}(\lvert t_0 \rvert)\right]$ ≈ 0.0226

**For a *t*-test with unknown and unequal variances for difference in populations, we must calculate the degrees of freedom differently

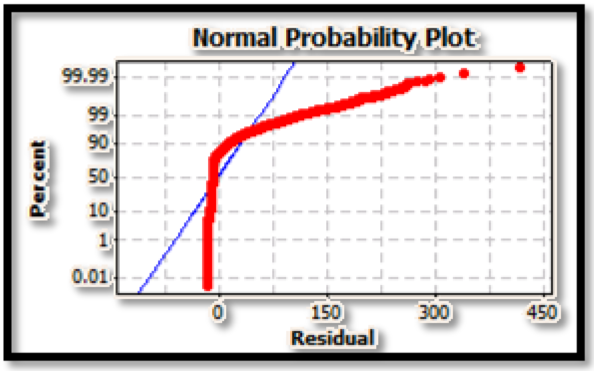
➔ Reject H0

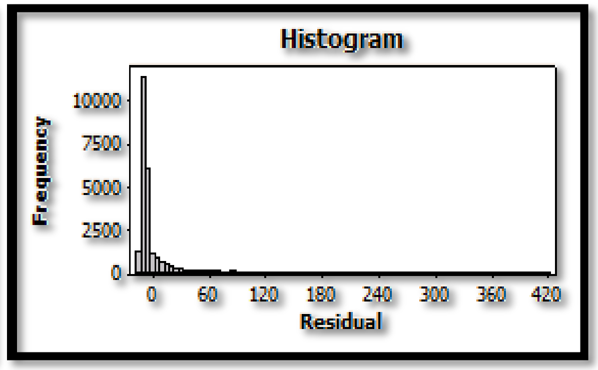
The results indicate that AirTran's average arrival delay is less than Delta's. This includes early-arriving flights. Interestingly, despite having average departure delay's greater than zero, both airlines still tend to arrive at their destinations before the scheduled time. A plot of the confidence intervals for difference in mean delay times is shown below; neither of the intervals encompasses zero, indicating 95% confidence that there is a difference between airlines.

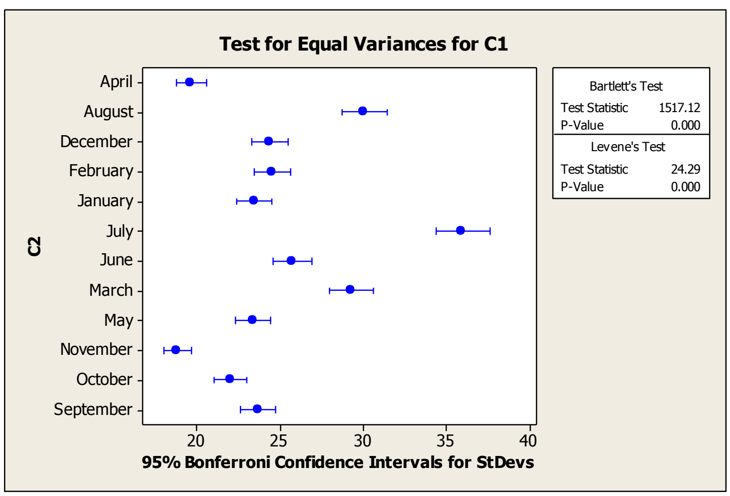


<u>Question V:</u> What time of the year are departure delays most likely to occur?

When analyzing departure delay throughout the year, an approach similar to the previous analysis was used. For each month a random sample of 2000 points was selected. For departure delays the tardiness of sample flights was capture by assuming an early departure to have zero delay. The data used for this case was not suitable for analysis of variance, since variance was not consistent among the samples from each month and the data did not follow a normal distribution, having a p-value = 0 for the residual normality test, which is shown below along with an histogram of the distribution, and the test for equal variances of the samples used.

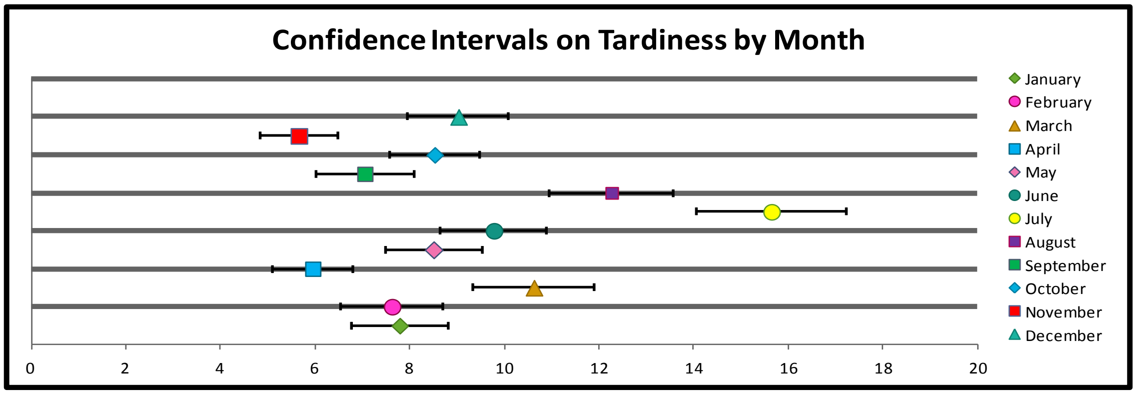






Since the data was not suitable for ANOVA, a Kruskal-Wallis test on the medians was performed. The resulting p-value for the test was zero, leading to the conclusion that the medians of all twelve samples are not equal.

As an alternative to the two tests previously mentioned a t-test with confidence interval of 95% was built for each of the twelve samples independently. The comparison of the means of tardiness for each month is shown below:



This graph shows that for the population that is being considered, the tardiness is greater in the month of July.

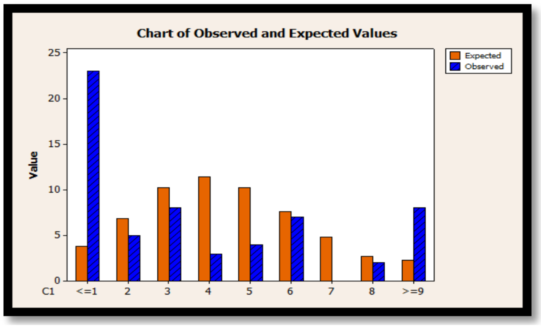
<u>Question VI</u>: Does the number of cancelled flights in a given day follow a Poisson distribution?

When considering the rich data sample of flights available for this analysis, cancellation of flights is an interesting aspect to be analyzed. A goodness of fit test was performed on the sample in an effort to determine whether the number of cancelled flights in a given day followed a Poisson distribution. Considering the definition of a Poisson distribution as of being a probability distribution that models the number of events in a given unit, this test seemed appropriate given the discrete nature of the data. Sixty random samples of number of cancelled flights in a given day were collected for this analysis. The hypotheses for this analysis were the following:

For this analysis a test statistic based on the difference between observed and expected frequencies was used:

Since it was assumed under H_0 that the distribution of number of cancelled flights is Poisson, $n(\lambda)$, the maximum likelihood estimator of λ, $\bar{X}$ was used to determine the estimation of λ. The expected frequencies for a Poisson distribution could then be calculated by intervals corresponding to the number of cancelled flights per day that were observed in the sample. The expected frequency, E_i, was calculated by multiplying the number of samples, n, by the Poisson probability of a given interval, $P(X)$, which is defined as:

Using Minitab, the expected frequencies are calculated and compared with the observed ones:



Intervals	Observed Frequecies	Expected Frequencies
<=1	23	3.8028
2	5	6.9554
3	8	10.376
4	3	11.6092
5	5	10.3912
6	7	7.7508
7	0	4.9555
8	2	2.7722
>=9	8	2.3869

Assuming a significance level of 5% and the estimation of a single parameter, the rejection condition for this test is:

where k is the number of intervals in the sample and is the number of estimated parameters. In Minitab H_0 was found to be 125.631 and 0. This result shows that the null hypothesis is rejected and the number of cancelled flights in a day does not follow a Poisson distribution.

<u>Question VII</u>: Which carrier has a lower proportion of cancelled flights, Delta or AirTran?

Another interesting statistic to make an inference upon is the proportion of cancelled flights of a given carrier. Even though the data set contains data of all airliners operating at ATL, only Delta, AirTran and Express Jet carriers were considered. A hypothesis test was performed to determine whether the proportion of cancelled flight from the Delta carrier was lower than the proportion of the AirTran carrier. The hypotheses and the rejection condition are shown below:

Rejection Region:

Since these populations had different sizes, a sampling rate of 0.5 was applied to the data. In this analysis, it is assumed that these populations don't have equal variances. The sample sizes and proportions are displayed in the table below:

Sample	X	N	Sample p
Delta	297	100199	0.002964
AirTran	129	30506	0.004229

Since the samples are large, the use of the Central Limit Theorem allows for the use of a standardized test statistic:

where

Using Minitab, the test statistic is found to be $Z_0 = -3.09$ and $p = 0.001$. With this result there's evidence that the proportion of cancelled flights of the Delta carrier is smaller than the AirTran carrier.

Conclusion:

The results of these studies yielded a number of interesting results that can be valuable in planning future air travel. In general, the average flight departure delay from Atlanta in 2012 was slightly greater than zero, although average arrival times still adhered to schedule. This illustrates an interesting trend that airlines are able to make up lost time along their routes, an ability that they certainly account for in flight planning. With respect to the two major air carriers from ATL, AirTran had lower average delays than Delta but a greater percentage of cancellation. Considering time of year and day of the week, July showed the worst travel delays, suggesting that summer vacation travel may play a significant role in increasing traffic and contributing to delays. Also, Saturday flights were the most on-time, which could be due to the tendency for weekend travel to occur on Friday and Sunday evening. Delays also tend to be minimal during the early morning and increase throughout the day. This may be due to the accumulation of traffic as well as early delays propagating into later flights.

The limitations of these results are that the data only came from the year 2012 and from one particular airport. Flight trends may be different for other years and could be affected by serious weather, economic climate, politics, and other unpredictable events. While the results are relevant to travelers based in Atlanta, the conclusions do not necessarily hold for other major hubs and are even more likely to differ for small, regional airports. Future investigations would significantly augment this study by expanding the scope to include a variety of airport locations and sizes, a period of several years, and more comparisons between major airlines and smaller carriers. The public source used to compile the data for this report contains all of the necessary information to perform these additional types of studies. Overall, a variety of statistical methods from this course were applied to glean some very interesting and relevant insights into the tardiness of airlines at our local Atlanta International Airport.

References:

1) Dataset Source: http://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=236&DB_Short_Name=On-Time
2) http://www.rita.dot.gov/bts/help/aviation/html/understanding.html

Appendix A: Descriptive statistics used in calculating the confidence intervals for causes in arrival delays

Delay Cause	# Data Points (n)	Mean ($\bar{X}$)	Standard Deviation (S)	Confidence Interval
Carrier Delay	1000	22.69	38.25	[20.32, 25.06]
Weather Delay	1000	1.48	11.67	[0.75, 2.20]
NAS Delay	1000	10.38	21.51	[9.05, 11.72]
Security Delay	1000	0.15	3.10	[-0.04, 0.34]
Late Aircraft Delay	1000	18.36	33.41	[16.29, 20.43]

Appendix B: Descriptive statistics used in calculating the confidence intervals for time of day

Delay Cause	# Data Points (n)	Mean ($\bar{X}$)	Standard Deviation (S)	Confidence Interval
Morning	1000	-7.303	33.393	[-9.37, -5.23]
Midday	1000	-2.145	24.443	[-3.66, -0.63]
Evening	1000	3.517	35.395	[1.32, 5.71]
Late-Night	1000	12.063	45.069	[9.27, 14.86]