Customer Satisfaction Rates for Four American Airlines from 1995 to 2018

David Fudge

August 25th, 2019

Overview

Every year for various airlines in the United States, the American Customer Satisfaction Index (ACSI) measures the overall customer satisfaction rates as a percentage. Data from this source were analyzed in R for four airline companies based in the United States for the years 1995-2018 by performing ANOVA. In addition, pairwise comparisons and tests of the assumptions were conducted. The primary conclusion drawn from the analysis was that Southwest Airlines performed very well in contrast to American, Delta, and United Airlines. Solutions proposed for the latter three airlines include: deals, decreasing flight delays, airplane maintenance, and effective customer service.

Pre-Diagnostics

Observe Figure 1 below, where the data is displayed by overlapping time series plots.

Figure 1: Overlapping time series plots for four airline companies.

Based on the data visualized above, consider various summary statistics for the considered airline companies.

Table 1: Summary Statistics

		N	Mean	Std. Dev.	Min	Median	Max
Company	American	24	65.54	4.33	60	64	76
	Delta	24	66.42	4.79	56	66.5	76
	Southwest	24	76.42	3.36	70	76	81
	United	24	62.75	4.1	56	62	70

Based on Table 1, the airlines American, Delta, and United have had less favorable customer satisfaction rates than Southwest Airlines over twenty-four year period in terms of both the median and average rates. Also, Delta Airlines has fluctuated more on average than the other airline companies since both the standard deviation and the range are relatively large.

The data in question is time series in nature. As such, it is necessary to check for autocorrelation which, if proven significant, drastically changes the entire analysis hereafter. To check autocorrelation for any of the four airlines, the use of ACF (autocorrelation function) plots as seen below in Figure 2 are helpful.

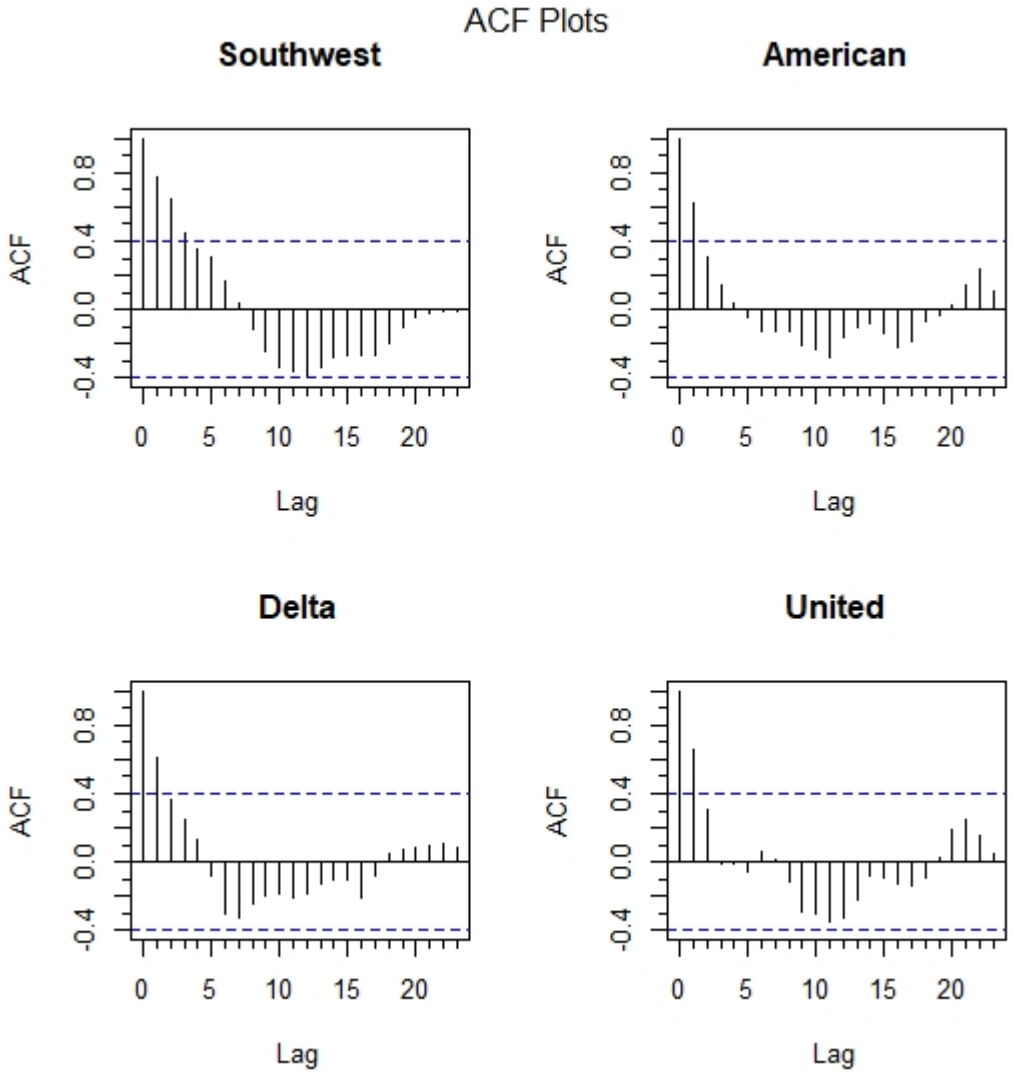

Figure 2: ACF plots by airline company.

For any of the four airline companies, the correlation coefficient measuring the correlation between the time series itself and the series once-lagged is above the benchmark of 0.4. In other words, for a given time series, the overall satisfaction rate for a year is somewhat correlated with the satisfaction rate the prior year. To be rid of time dependencies, ten "random" years (pseudo-randomly generated in R) without replacement were chosen for each time series. Side-by-side boxplots were generated below in Figure 3.

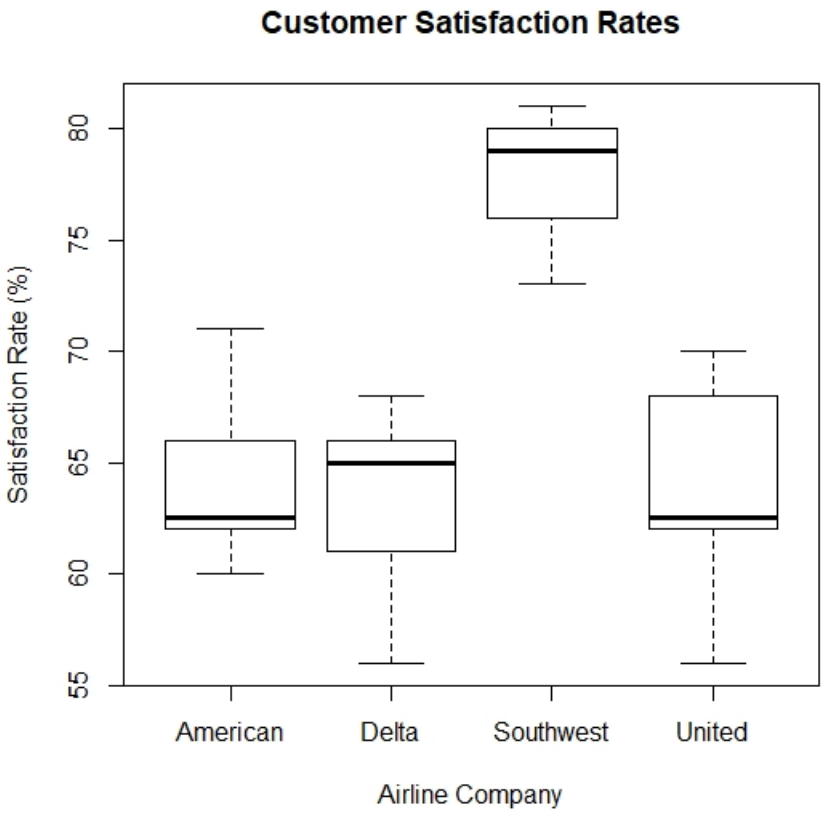

Figure 3: Side-by-side boxplots of customer satisfaction rates by airline company.

Figure 3 identifies that Southwest Airlines' rates are likely different from all other considered airlines. Among American, Delta, and United Airlines, any difference in satisfaction rates is difficult to determine.

Analysis of Variance

For the purposes of ANOVA (analysis of variance), the null hypothesis assumes that all airlines are the same with respect to customer satisfaction rates. Since the years were chosen at random, the year cannot be included as a factor that affects satisfaction rates. However, the airlines themselves are a likely factor. Because the airline companies were chosen at random, the design of this experiment is a single-factor fixed-effects design. For consistency on significance, the significance level α will be set to 0.01. Table 2 below provides the ANOVA table, as computed in R.

Table 2: ANOVA Table

	SS	df	MS	F	p-value
Airline	1569.9	3	523.3	37.74	3.29×10^{-11}
Residuals	499.2	36	13.9		

The p-value (3.29×10^{-11}) for the factor describing the airline is below $\alpha = 0.01$. Then there exists a significant difference in customer satisfaction rates between at least two of the airline companies considered. Next, before any pairwise comparisons between airlines can be confidently reported, two underlying assumptions (normality of the residuals and homogeneity of variance) must be met.

Test of Normality

One critical assumption for this ANOVA design is that the residuals are normally distributed. Figure 4 below gives the qq-normal plot of the residuals of the experiment.

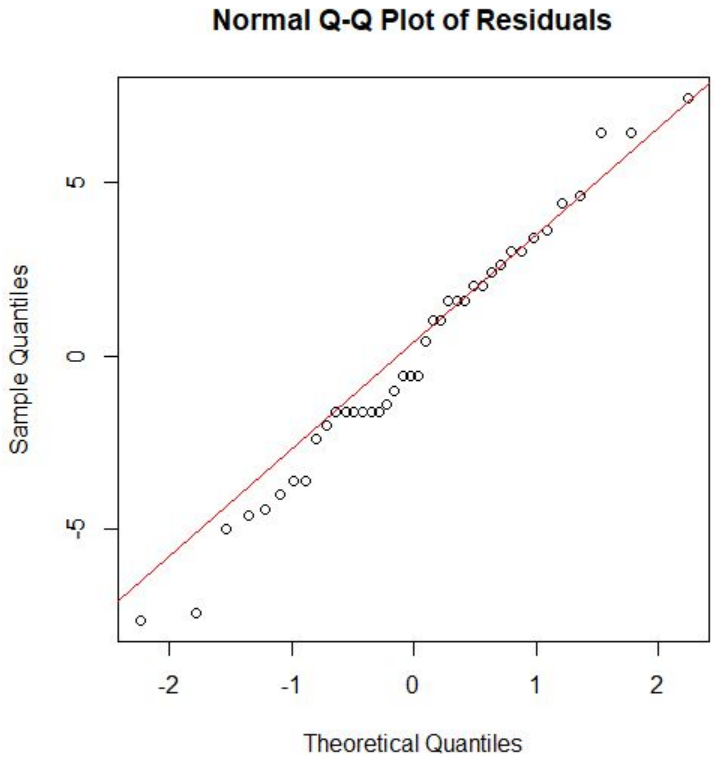

Figure 4: The qq-normal plot of the residuals.

Overall, there appears to be a linear trend in the qq-normal plot, although the lower end of the data is wave-like. A robust hypothesis test for normality, as an alternative to the qq-plot

in Figure 4, is the Shapiro-Wilk test, whose null hypothesis assumes normality. Applying the test to the residuals, the test statistic is $W = 0.98303$ and the p-value associated is 0.7997. The p-value is much larger than the significance level $\alpha = 0.01$, so the assumption of normality should not be rejected. Furthermore, the results from both the qq-normal plot and the Shapiro-Wilk test suggest that the assumption that the residuals are normally distributed is met with confidence.

Test of Constant Variance

The other assumption imposed in this design of experiment is homogeneity of variance. A visual indicator of constant (or nonconstant) variance is a residuals vs. fitted values plot in Figure 5 below.

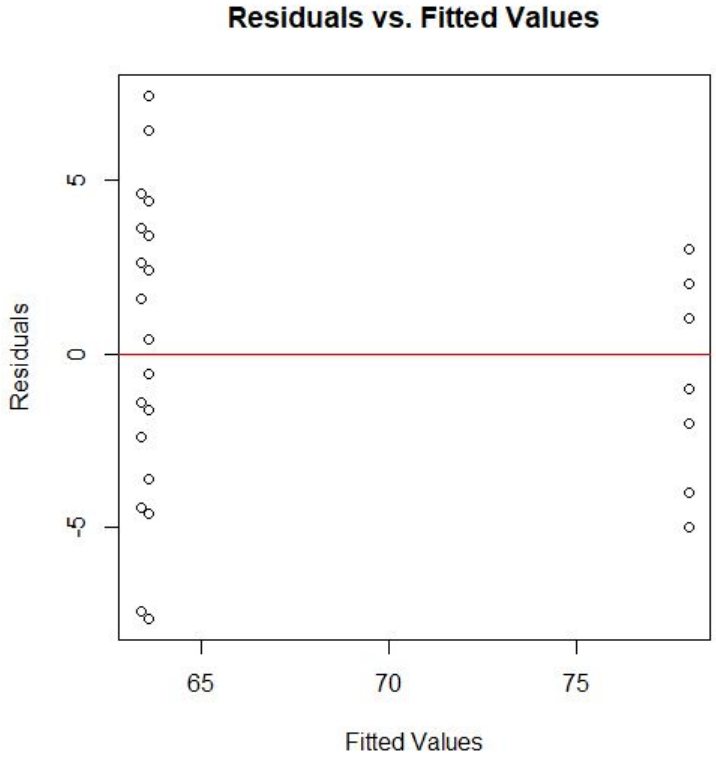

Figure 5: The residuals vs. fitted values plot.

The residuals in the plot express a very slight concern in that there is a very slight inward trend or funnel. With this concern, it is necessary to use hypothesis test. One such test, where the null hypothesis assumes constant variance, is Modified Levene's test. For Modified Levene's test in this analysis, the F-statistic is 0.3892 and the associated p-value is

0.7615. Since the p-value is much greater than $\alpha = 0.01$, it would be unwise to reject the null hypothesis. Moreover, the results from Modified Levene's test imply that the assumption of constant variance is satisfied. Because the p-value is so large in comparison to α, the numerical argument for keeping the assumption seems to outweigh the visual evidence to the contrary.

Pairwise Comparisons

Earlier it was determined from the ANOVA table that at least one airline company differs from another. Also, since the assumptions of ANOVA appear to be satisfied, the results appear trustworthy. Therefore, pairwise comparisons of means can now be reliably made between airline companies. The test used in this experiment was Tukey's test, where the null hypothesis assumes that the mean satisfaciton of one airline company is equal to the mean of another. Because four companies were considered in ANOVA, a total of six pairwise comparisons can be made. The results of Tukey's Test with an overall 95% confidence level can be seen below in Table 3.

Table 3: Tukey's Test for Pairwise Comparisons of Means

			difference	lower est.	upper est.	adj. p-value
American	-	Delta	-0.2	-4.69	4.29	0.9993687
American	-	Southwest	14.4	9.91	18.89	~ 0
American	-	United	0.0	-4.49	4.49	~ 1
Delta	-	Southwest	14.6	10.11	19.09	~ 0
Delta	-	United	0.2	-4.29	4.69	0.9993687
Southwest	-	United	-14.4	-18.89	-9.91	~ 0

At the significance level $\alpha = 0.01$, only the p-values in Table 3 associated with Southwest Airlines are less than α. This implies that the mean customer satisfaction rate for Southwest Airlines is significantly different to the rates of the airlines American, Delta, and United. These results numerically verify the apparent differences in airlines from both Figures 1 and 3. There is no indication of a significant difference between the satisfaction rates for American, Delta, and United Airlines.

Kruskal-Wallis Test

Most of the results above rely on the assumptions of the single-factor fixed effects ANOVA. If there is any doubt in validity, a quick nonparametric alternative to ANOVA,

called the Kruskal-Wallis Test, can be performed. Unfortunately, Tukey's Test does not come in quite as handy as before because Tukey's Test relies on ANOVA.

The chi-squared statistic associated with the Kruskal-Wallis Test is 22.109. The p-value is 6.19×10^{-5}. Similar to the results of ANOVA, with significance level $\alpha = 0.01$, the p-value is smaller than α. Therefore, there is likely some difference in customer satisfaction rates between the airline companies chosen for the experiment.

Conclusion

The goal of this experiment was to determine any potential differences in mean customer satisfaction rates between four U.S.-based airline companies: American, Delta, Southwest, and United Airlines. Accounting for any possible difference in rates, analysis of variance was performed by taking a random sample. The conclusion made from the ANOVA is that there is very likely a difference between at least two airline companies. Overall, the model behind the design proved to be reliable for this experiment. Keeping that in mind, pairwise comparisons were made, and most companies had a statistically significant difference in satisfaction rates.

Very clearly, Southwest Airlines has had the best satisfaction rates between 1995 and 2018. Let alone that Figure 1 shows that Southwest Airlines hovers above the other three airlines for any given year, the difference is statistically significant according to ANOVA and Tukey's Test.

The satisfaction rate for a company reflects the willingness for a customer to return for another (or the same) product. In terms of airlines, the cheapest airline will be most often chosen for travel regardless of perceived quality, unless price differences are relatively small. Factors such as how long a flight is delayed or how responsive and cooperative customer service is will affect how satisfied a customer is. In the worst case scenario, the cost to fly is high, the flight is significantly delayed, and the conditions in the plane are bad. As such, American, Delta, and United Airlines should examine the business models of Southwest Airlines. Solutions to increase low airline satisfaction rates may include: providing deals for groups of customers, preventing airplane delays as much as possible, maintaining airplane features such as the fans and lights above seats, and ensuring an employee policy of responsive customer service.

References

American Customer Satisfaction Index. *Benchmarks by Industry*. Available from https://www.theacsi.org/index.php?option=com_content&view=article&id=149&catid=14&Itemid=212&i=Airlines