REPORT ON FAA Project

Name: Gopalakrishnan Kalarikovilagam Subramanian UID: M12821535

Summary:

The goal of the project is to develop a model for the landing distance as a function of the variables speed_air, speed_ground, height, pitch, duration, no_psg and aircraft. Two data sets were first merged into a single one after removing missing rows.The data set was first cleaned and duplicate rows were removed. The missing and abnormal values were found. The rows containing the abnormal values were removed. The variables were plotted with respect to the distance once the data has been cleaned. The relationships were guessed based on the plots. Then, the correlation test was carried between all two pairs of variables. There was found to strong positive correlation between distance and speed_air/speed_ground. There was also correlation between distance and pitch, distance and height. Strong Internal positive correlation was found between speed_air and speed_ground. Similar analysis was done after splitting the data set for Boeing and Airbus aircrafts. For Boeing there was positive correlation between distance and speed_air/speed_ground. For airbus there was correlation between distance and speed_air/speed_ground and height. In the next step the variables where correlation was found was taken to do the modeling. The R-square was used as an indicator of how strong the model is. For the aircraft data set there was a good fit between distance and speed_ground, height and pitch. For the Boeing aircraft a model was formed between distance and speed_ground. Similarly, for the Air-

bus aircraft a model was formed between distance and speed_ground and height. The residuals for all the models passed the T-Test for zero mean value.

1. How many observations (flights) do you use to fit your final model? If not all 950 flights, why?

All the observations were not used because the duplicate observations in the data set was removed.Then the abnormal data points were removed.

2. What factors and how they impact the landing distance of a flight?

The landing distance was affected by speed_ground, speed_air, height and pitch. The speed_ground and speed_air had a positive correlation with landing distance. The sign of the correlation cannot be ascertained for pitch and height.

3. Is there any difference between the two makes Boeing and Airbus?

Yes, there is a difference between the makes of Boeing and Airbus. For the Boeing aircraft there is a correlation between distance and speed_ground/speed_air alone. However, for Airbus there is a correlation between distance and speed_ground/ speed_air and height. The Boeing aircraft has a higher mean and standard deviation of landing distance compared to Airbus. Also the observations where the landing distance is more than 6000m has also happened for Boeing aircraft.

CHAPTER 1: DATA Preparation

Goal:

The goal of the project is:

There are two excel files FAA1 and FAA2. FAA2 contains 151 rows and FAA1 contains 801 rows. The variables in FAA1 are aircraft, duration, no_psg, speed_ground, speed_air, height, pitch and distance. The variables in FAA2 are aircraft, no_psg, speed_ground, speed_air, height, pitch and distance. Therefore, the variable duration is missing in FAA2. The two excel sheet data are imported in SAS environment. The rows in which all the variables are missing is removed from both the imported data sets. Then the two excel-sheet data are merged using method 1 (concatenating data sets). The code for these two is given below:

1. Combining data sets from different sources;
2. Performing the completeness check of each variable – examine if missing values are present;
3. Performing the validity check of each variable – examine if abnormal values are present;
4. Cleaning the data based on the results of Steps 2 and 3;
5. Summarizing the distribution of each variable

CODE:

```
/* Generated Code (IMPORT) */

/* Source File: FAA1.xls */

/* Source Path: /folders/myfolders/Stat_Computing */

/* Code generated on: 9/13/18, 5:45 PM */
```

```
%web_drop_table(WORK.IMPORT);

FILENAME REFFILE '/folders/myfolders/Stat_Computing/FAA1.xls';

PROC IMPORT DATAFILE=REFFILE

   DBMS=XLS

   OUT=WORK.aircraft1;

   GETNAMES=YES;

RUN;

/* To remove all observations with all variables empty from the first excel file data*/

data aircraft1_mod;

set aircraft1;

if compress(cats(of _all_),'.')=' ' then delete;

run;

PROC CONTENTS DATA=WORK.aircraft1_mod; RUN;

/* Import the second excel file */

%web_open_table(WORK.IMPORT);

%web_drop_table(WORK.IMPORT);

FILENAME REFFILE '/folders/myfolders/Stat_Computing/FAA2.xls';

PROC IMPORT DATAFILE=REFFILE

   DBMS=XLS
```

```
   OUT=WORK.aircraft2;
   GETNAMES=YES;
RUN;
/* To remove all observations with all variables empty from the second excel file data */
data aircraft2_mod;
set aircraft2;
if compress(cats(of _all_),'.')=' ' then delete;
run;
proc means data =aircraft2_mod;
run;
PROC CONTENTS DATA=WORK.aircraft2_mod; RUN;
%web_open_table(WORK.IMPORT)
/*merge two data sets */
data aircraft;
set aircraft1_mod aircraft2_mod;
run;
/*To remove duplicate observations from a dataset */
proc sort data= aircraft nodupkey;
by distance speed_air speed_ground height pitch ; Run;
```

/* To tabulate the number of missing values in different variables */

```
proc means data=aircraft NMISS N; run;

data aircraft_boeing;

set aircraft;

keep;

where aircraft='boeing';

run;

proc print data=aircraft_boeing;

run;
```

/* Count the number of values with abnormal duration */

```
data aircraft_copy;

set aircraft_boeing;

label duration = Number of abnormal duration;

where duration < 40 and duration ^=.;

run;

proc means data =aircraft_copy n;

var duration;

run;
```

/* Mark all obervations with missing values as YES under a new variable missing_values*/

```
data aircraft_proper;

set aircraft_boeing;

if (duration =. OR speed_ground =. OR speed_air =. OR

height =. OR pitch =. OR distance =.) then missing_values ='YES';

else missing_values = 'NO';

run;

proc means data=aircraft_proper NMISS N; run;

/* Mark all obervations with abnormal values as YES under a new variable abnormal_values*/

data aircraft_proper2;

set aircraft_proper;

if ((duration < 40 and duration~=.)or((speed_ground < 30 OR speed_ground > 140) and speed_ground
        ~=.)

OR ((speed_air < 30 OR speed_air > 140) and speed_air ~=.) OR (height < 6 and height ~=.) or
        (distance>=6000 and distance~=.))

then abnormal_values ='YES';

else abnormal_values = 'NO';

run;

/* Histogram plot of duration variable */

ods noproctitle;

ods graphics / imagemap=on;
```

```
/* Exploring Data */

proc univariate data=WORK.AIRCRAFT_PROPER2;

   ods select Histogram;

   var distance;

   histogram distance;

   inset n mean median std var skewness min max / position=ne;

run;

proc stdize data=aircraft_proper2 reponly method=mean out=aircraft_proper3;

var duration speed_ground speed_air height pitch distance;

run; */

/*removal of rows with abnormal values */

data aircraft_cleaned;

set aircraft_proper2;          /* open data set for edit */

if (abnormal_values = "YES") then delete;

run;

proc means data=aircraft_cleaned NMISS N; run;

proc print data = aircraft_cleaned;

run;
```

RESULTS:

Table1: The list of variables in the data set

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	Label
1	aircraft	Char	12	$12.	$12.	aircraft
8	distance	Num	8	BEST12.		distance
2	duration	Num	8	BEST12.		duration
6	height	Num	8	BEST12.		height
3	no_pasg	Num	8	BEST12.		no_pasg
7	pitch	Num	8	BEST12.		pitch
5	speed_air	Num	8	BEST12.		speed_air
4	speed_ground	Num	8	BEST12.		speed_ground

Table 2: showing that 50 rows were totally missing in the second excel file.

Variable	Label	N	N Miss
no_pasg	no_pasg	150	50
speed_ground	speed_ground	150	50
speed_air	speed_air	39	161
height	height	150	50
pitch	pitch	150	50
distance	distance	150	50

Variable	Label	N	N Miss
no_pasg	no_pasg	150	0
speed_ground	speed_ground	150	0
speed_air	speed_air	39	111
height	height	150	0
pitch	pitch	150	0
distance	distance	150	0

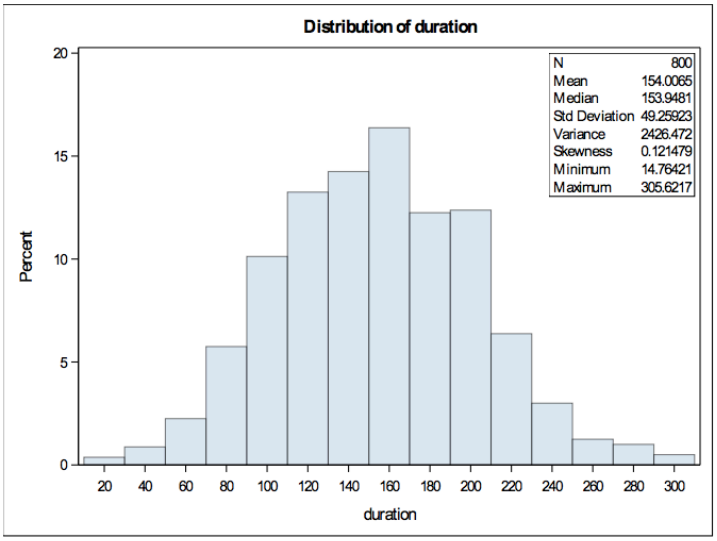
Table 3: Showing the number of each variables and the number of missing values:

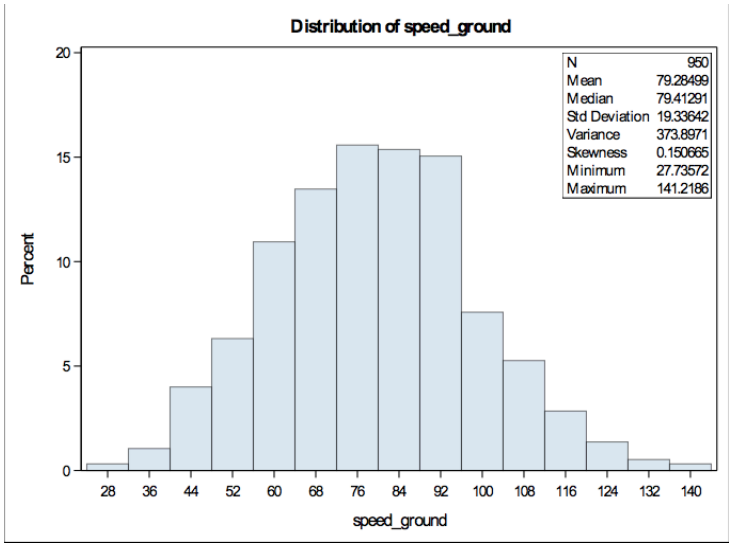
Variable	Label	N Miss	N
duration	duration	150	800
no_pasg	no_pasg	0	950
speed_ground	speed_ground	0	950
speed_air	speed_air	711	239
height	height	0	950
pitch	pitch	0	950
distance	distance	0	950

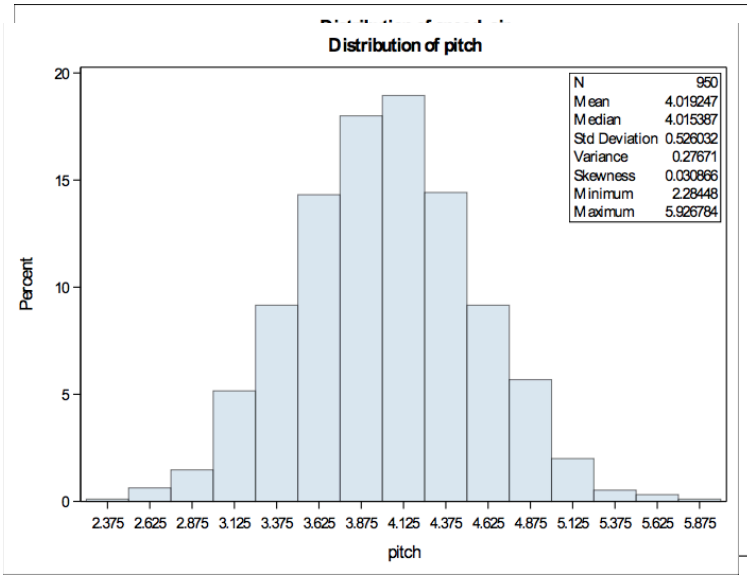
Table 5: showing a snap shot of the new variables in the table. It shows weather there is a missing value or abnormal value in the row or not.

Obs	aircraft	duration	no_pasg	speed_ground	speed_air	height	pitch	distance	missing_values	abnormal_values
1	airbus	153.94809754	46	40.801786477	100.891677	24.400127629	3.9682093233	620.09051196	YES	NO
2	airbus	153.94809754	46	104.07757658	103.40921036	19.7157721	4.1043931104	2494.8046454	YES	NO
3	airbus	153.94809754	48	61.570704648	100.891677	21.785707448	4.3511947442	560.53392302	YES	NO
4	airbus	153.94809754	50	84.219908138	100.891677	32.542946798	3.318828622	1485.4400456	YES	NO
5	airbus	153.94809754	51	62.484050366	100.891677	26.53804471	3.8228939729	749.48028928	YES	NO
6	airbus	153.94809754	51	83.630692914	100.891677	23.302265488	4.5566399591	1460.4181796	YES	NO
7	airbus	153.94809754	52	72.036625004	100.891677	24.740341243	3.6279838777	648.02156805	YES	NO
8	airbus	153.94809754	52	73.761115944	100.891677	9.688307724	3.3585464091	554.16098701	YES	NO
9	airbus	153.94809754	52	89.577029476	100.891677	35.463228123	3.834651479	1390.8995718	YES	NO
10	airbus	153.94809754	54	50.903105868	100.891677	35.729484049	4.5440403076	597.98554514	YES	NO
11	airbus	153.94809754	54	67.456935552	100.891677	41.334169856	3.8581993926	877.06227359	YES	NO

Histogram plot of duration, speed_ground, speed_air, height, pitch and distance. It can be seen







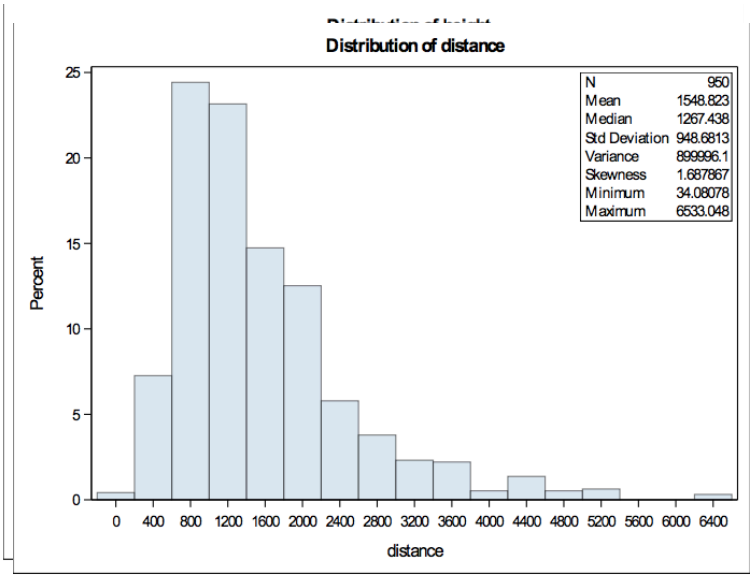


Table 7: Showing the coefficient of variation for all the variables in the data set.

Variable	Label	Coeff of Variation
duration	duration	31.9851574
no_pasg	no_pasg	12.4490507
speed_ground	speed_ground	24.3884962
speed_air	speed_air	10.2237258
height	height	34.3715976
pitch	pitch	13.0878300
distance	distance	61.2517438

OBSERVATIONS:

It can seen from the tables that there are lot of missing observations in the duration and speed air variables. The duration variable is missing from the second excel file.

1. While importing the excel files it was found that there were 50 rows missing from the second data set and they have been removed. This is shown in the table 2

2. The minimum maximum, standard deviations, means, medians of the combined data set is shown above in the histogram plot of each variable. There are no absurd values except for the height variable where we have a negative distance.

3. The mean and median are pretty close to each other for most of the variables except for speed_air and distance.

4. The missing values were found in the two variables duration and speed_air.

5. The missing values have been kept as it is as no other information was provided about the missing values.

6. The rules for handling the abnormal observations were also not provided. Hence, the rows containing the abnormal values were deleted.

7. From the histograms shown above it can be seen that most of the variables follow normal distribution except speed of air and distance. The skewness value is positive for both of these variables.

8. The coefficient of variation is highest for distance followed by height, duration and ground speed. This shows the dispersion is highest for the distance variable which can be seen from the histogram plot too.

Variables	Missing values	Abnormal values	% missing values	% Abnormal values	%missing after removal
Speed_ground	0	12		1.4	
Speed_air	642	1	75.5	0.1	74.59
Height	0	10		1.1	
Pitch	0	NA			
Duration	50	5	5.8	0.6	15.8
No_psg	0	NA			
Distance	0	2		0.2	

CONCLUSION:

All the mentioned targets in goals were completed. The two excel files were imported as data sets in SAS. They were checked for missing rows and cleaned. The two data sets were combined together and checked if there was any observation repetition. The combined data set was checked for missing values and abnormal values. The missing values were kept as it is and the abnormal values were deleted from the data set. A distribution plot was made for all the variables and observations were drawn from it.

CHAPTER 2: DATA VISUALISATION

GOAL:

The goal of the chapter is :

1. To take the cleaned data and plot the graphs between the output variable landing distance and all the input variables.
2. Visualise how the relationship is between the input variables and all the output variables.
3. Make a correlation analysis between all the variables.
4. From the correlation values determine which all variables have a strong correlation with the output variable and check if any input variables have strong internal correlation.

5. Determine which all variables need to be chosen for the modelling part in the next chapter.

CODE:

The code below is for making a X-Y plot between different variables

```
proc plot data=aircraft_cleaned;
   plot distance*speed_ground = '.';
   plot distance*duration = '.';
   plot distance*speed_air = '.';
   plot distance*height = '.';
   plot distance*pitch = '.';
```

```
run;
```

The code below is for making a correlation analysis between different pairs of variables:

```
proc corr data=aircraft_cleaned;
var distance speed_ground speed_air height pitch duration;
run;
```

The code below is for creating a new data set consisting of only those rows where the aircraft is Boeing.

```
data aircraft_boeing;
set aircraft;
keep;
where aircraft='boeing';
run;
```

Similarly, the below code is for creating a new data set consisting of only those rows where the aircraft is Airbus.

```
data aircraft_airbus;
set aircraft;
keep;
where aircraft='airbus';
run;
```

The code below is for making an analysis of the variable 'landing distance' when the aircraft is Boeing or Airbus.

```
proc means data=aircraft_boeing;
run;
proc means data=aircraft_airbus;
run;
```

RESULTS:

The plot of distance versus speed ground is given belowFig

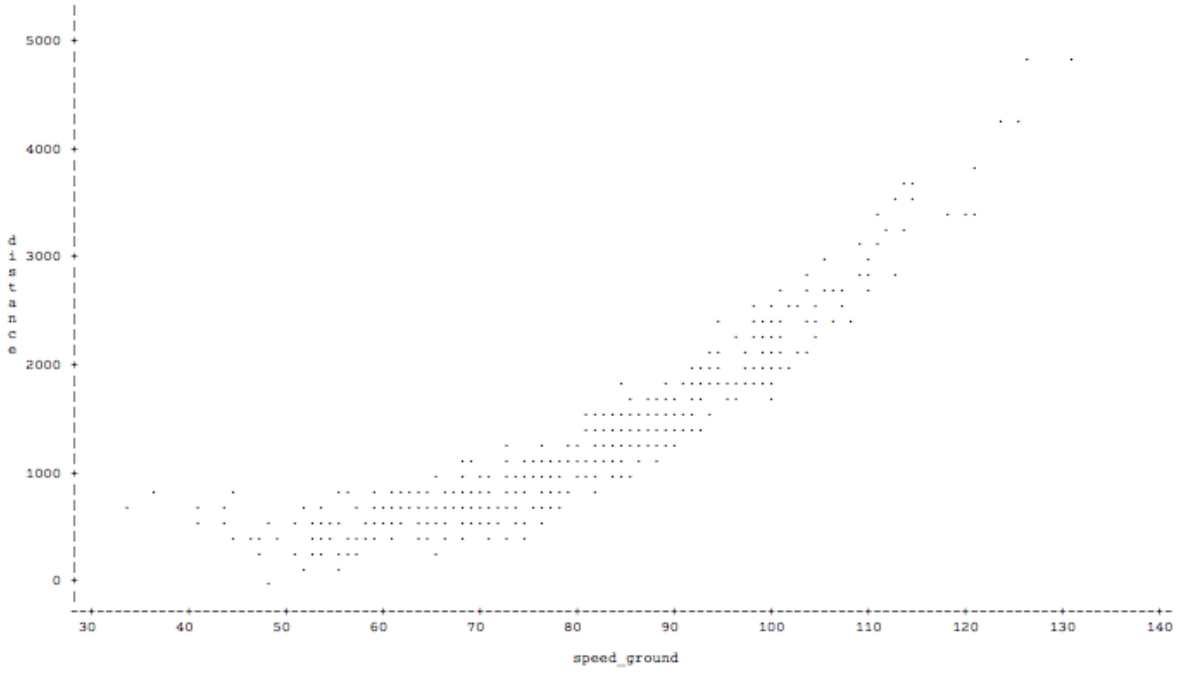


Fig 1: The plot of distance vs Speed_ground

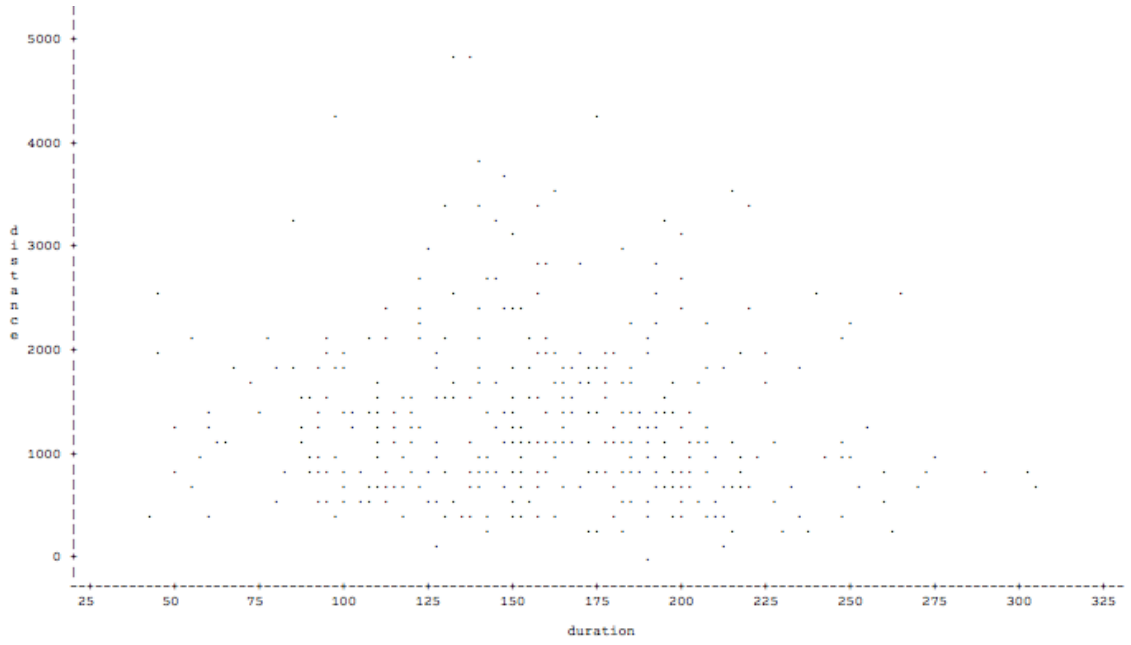


Fig 2: The plot of distance vs duration

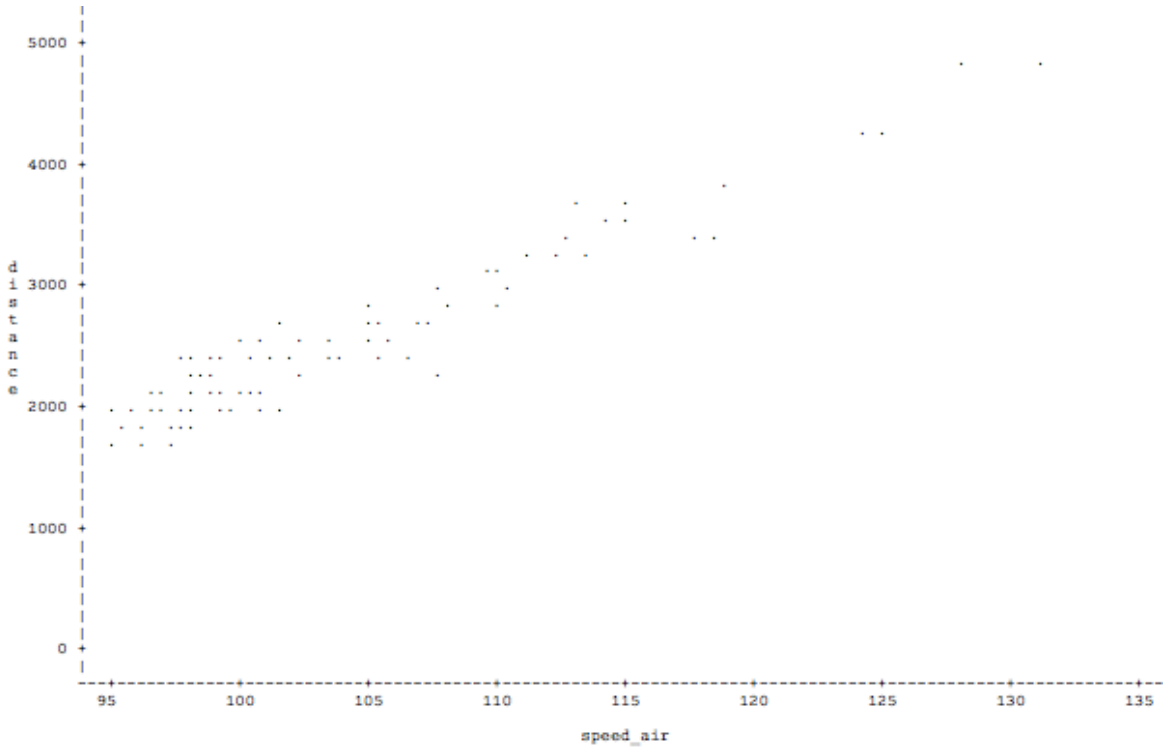


Fig 3: The plot of distance vs speed_air

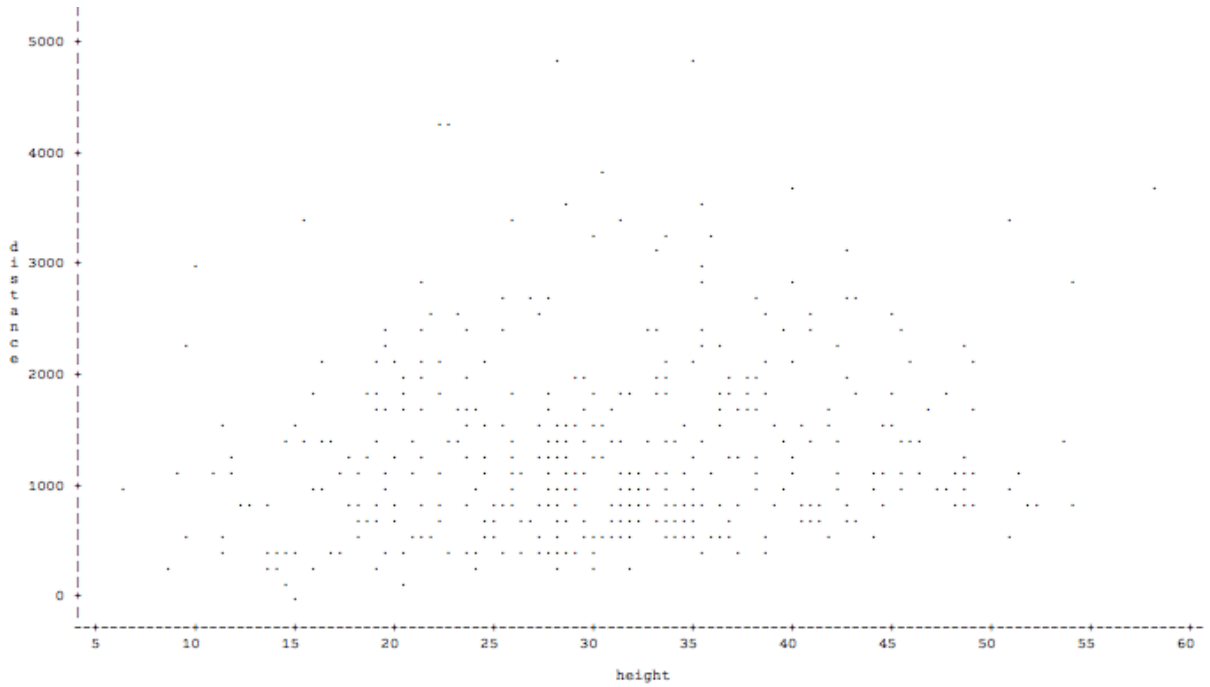


Fig 4: The plot of distance vs height

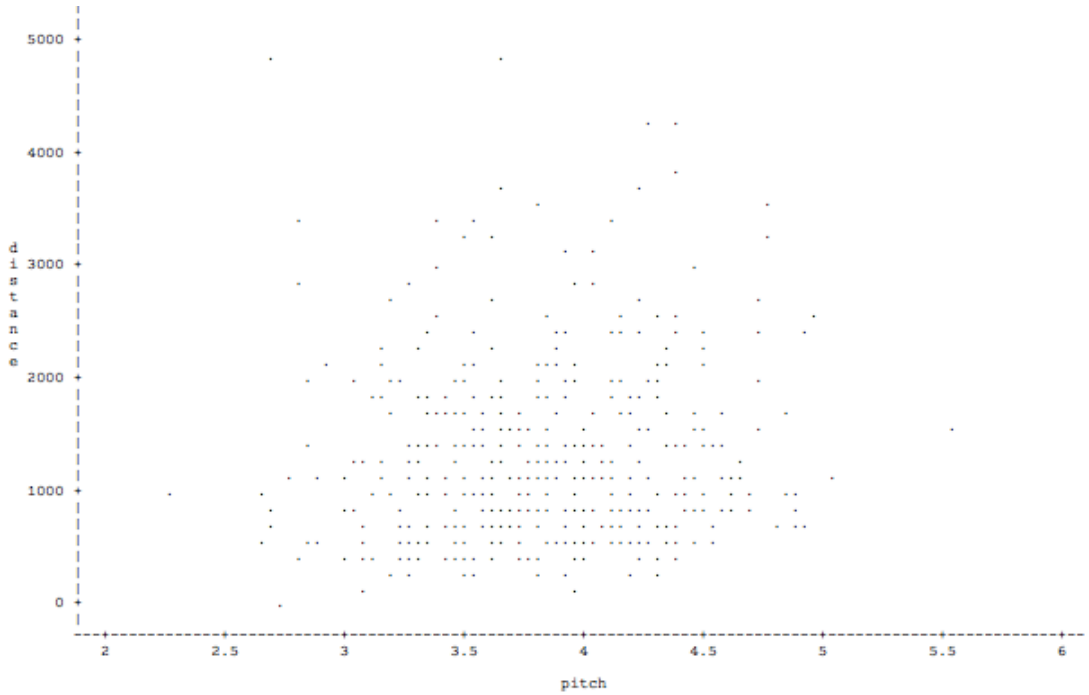


Fig 5: The plot of distance vs pitch

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations						
	distance	**speed_ground**	**speed_air**	**height**	**pitch**	**duration**
distance distance	1.00000 831	0.86624 <.0001 831	0.94210 <.0001 203	0.09941 0.0041 831	0.08703 0.0121 831	-0.05138 0.1514 781
speed_ground speed_ground	0.86624 <.0001 831	1.00000 831	0.98794 <.0001 203	-0.05761 0.0970 831	-0.03912 0.2599 831	-0.04897 0.1716 781
speed_air speed_air	0.94210 <.0001 203	0.98794 <.0001 203	1.00000 203	-0.07933 0.2606 203	-0.03927 0.5780 203	0.04454 0.5364 195
height height	0.09941 0.0041 831	-0.05761 0.0970 831	-0.07933 0.2606 203	1.00000 831	0.02298 0.5082 831	0.01112 0.7564 781
pitch pitch	0.08703 0.0121 831	-0.03912 0.2599 831	-0.03927 0.5780 203	0.02298 0.5082 831	1.00000 831	-0.04675 0.1918 781
duration duration	-0.05138 0.1514 781	-0.04897 0.1716 781	0.04454 0.5364 195	0.01112 0.7564 781	-0.04675 0.1918 781	1.00000 781

Table 1: Showing the correlation between the different variables

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
duration	duration	400	151.9031187	48.7011866	14.7642071	298.5223339
no_pasg	no_pasg	400	59.9425000	7.5834005	29.0000000	82.0000000
speed_ground	speed_ground	400	78.6118058	21.1999092	27.7357153	141.2186354
speed_air	speed_air	122	103.5054579	11.5689208	90.0028586	141.7249357
height	height	400	29.9468397	10.3387152	-3.5462524	59.9459639
pitch	pitch	400	4.2091735	0.4874715	2.9931514	5.9267842
distance	distance	400	1759.84	1012.25	371.2772609	6533.05

Table2: Basic statistical information of the variables for Boeing aircraft.

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
duration	duration	400	156.1099583	49.7830202	16.8934549	305.6217107
no_pasg	no_pasg	450	60.2466667	7.4174927	36.0000000	87.0000000
speed_ground	speed_ground	450	80.1994492	16.9206507	33.5741041	131.0351822
speed_air	speed_air	86	104.2123333	8.0924561	95.0113646	131.3379485
height	height	450	30.3196736	10.2505068	-3.3323880	58.2277997
pitch	pitch	450	3.8317436	0.5004493	2.2844801	5.5267842
distance	distance	450	1318.19	792.3479576	34.0807833	4896.29

Table 3: Basic statistical information of the variables for Airbus aircraft

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations					
	distance	**speed_ground**	**speed_air**	**height**	**pitch**
distance distance	1.00000 387	0.90050 <.0001 387	0.97760 <.0001 118	0.06920 0.1743 387	-0.06504 0.2017 387
speed_ground speed_ground	0.90050 <.0001 387	1.00000 387	0.99048 <.0001 118	-0.08263 0.1046 387	-0.04755 0.3509 387
speed_air speed_air	0.97760 <.0001 118	0.99048 <.0001 118	1.00000 118	-0.12922 0.1631 118	-0.02499 0.7882 118
height height	0.06920 0.1743 387	-0.08263 0.1046 387	-0.12922 0.1631 118	1.00000 387	0.00492 0.9232 387
pitch pitch	-0.06504 0.2017 387	-0.04755 0.3509 387	-0.02499 0.7882 118	0.00492 0.9232 387	1.00000 387

Table 4: Correlation between different variables for Boeing aircraft alone

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations					
	distance	**speed_ground**	**speed_air**	**height**	**pitch**
distance distance	1.00000 444	0.90520 <.0001 444	0.96411 <.0001 85	0.14494 0.0022 444	0.07330 0.1230 444
speed_ground speed_ground	0.90520 <.0001 444	1.00000 444	0.98169 <.0001 85	-0.03346 0.4819 444	-0.00493 0.9176 444
speed_air speed_air	0.96411 <.0001 85	0.98169 <.0001 85	1.00000 85	-0.00546 0.9604 85	0.00007 0.9995 85
height height	0.14494 0.0022 444	-0.03346 0.4819 444	-0.00546 0.9604 85	1.00000 444	0.05128 0.2809 444
pitch pitch	0.07330 0.1230 444	-0.00493 0.9176 444	0.00007 0.9995 85	0.05128 0.2809 444	1.00000 444

T

Table 5: Correlation between different variables for Airbus aircraft

OBSERVATIONS:

1. From the figure 1 and 3 it can be inferred that speed_ground and speed_air has a strong positive correlation with landing distance.
2. From Fig 2 ,4 and 5 it cannot be cleared about any relationship between the variables duration, height and pitch with landing distance.
3. The above inferences can be confirmed from the correlation table 1. It is seen from the table that there is a strong positive correlation between speed_ground, speed_air and landing distance.
4. Height and pitch are also correlated with landing distance, since the p value of the hypothesis test is small. However, the sign of the correlation cannot be ascertained as the Pearson correlation coefficient value is small.
5. Duration is not correlated with landing distance as the p value is quite large.
6. There is also a strong positive correlation between speed_air and speed_-ground. This can be ascertained by the p value which is very small and the Pearson correlation coefficient which is quite large.
7. From the basic statistical information for Boeing aircraft shown in Table 2 and Table 3 it is seen that the aircraft has crossed the prescribed distance of 6000m for Boeing aircraft but not for the Airbus aircraft.
8. From the correlation table between variables for Boeing aircraft alone as shown in Table4, there is a strong positive correlation between landing distance and speed_ground and speed_air. The other variables do not have a correlation with landing distance.
9. From the Table 5 it can be seen that the landing distance is correlated with speed_air, speed_ground and height for Airbus aircraft. There is a strong positive correlation between distance and speed_air and distance and speed_-ground.

CONCLUSIONS:

From the above observations it can be concluded that for the aircraft as a whole there is a correlation between distance variable and speed_air, speed_ground, height and pitch. Since, there is a correlation between speed_air and speed_ground it is preferred if we take speed_ground, since it has more observations. Therefore, the variables speed_ground, height and pitch will be used for the linear regression model of landing distance.

When we take Boeing aircraft alone, there is a correlation only between landing distance and speed_air/speed_ground. Again, as there is a correlation between speed_air and ground we are choosing speed_ground. In the case of airbus, there is a correlation between landing distance and speed_air/speed_ground and height.

The mean, standard deviation, minimum and maximum values of landing distance are higher for Boeing aircraft as compared to Airbus. The places where the aircraft has crossed the allowable distance of 6000m has also happened for Boeing aircraft.

CHAPTER 3: MODELLING

GOAL:

The goal of this chapter is to form a linear regression model between the observed values and the input variables. We are making a regression model for the entire aircraft data and also for Boeing and Airbus separately.

CODE:

Aircraft_cleaned is the cleaned data set that is obtained after removing the abnormal values from the dataset.

```
proc reg data=aircraft_cleaned;
model distance=speed_ground;
title regression analysis of the landing distance;
run;

proc reg data=aircraft_cleaned;
model distance=speed_ground height pitch;
title regression analysis of the landing distance;
run;

proc ttest data=diagnostics;
var residual;
run;
```

RESULTS:

Root MSE	276.59838	R-Square	0.8910
Dependent Mean	2787.37914	Adj R-Sq	0.8901
Coeff Var	9.92324		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-5442.44229	190.69213	-28.54	<.0001
speed_ground	speed_ground	1	-9.08063	11.86089	-0.77	0.4447
speed_air	speed_air	1	88.75128	11.99061	7.40	<.0001

Fig 1: Speed_ground/Speed_air vs distance

Root MSE	401.32240	R-Square	0.8061
Dependent Mean	1549.62400	Adj R-Sq	0.8055
Coeff Var	25.89805		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-3050.02669	126.48756	-24.11	<.0001
speed_ground	speed_ground	1	43.92893	0.72017	61.00	<.0001
height	height	1	13.44958	1.34514	10.00	<.0001
pitch	pitch	1	170.00672	25.37622	6.70	<.0001

Fig 2: Speed_ground,height,pitch vs Distance

Root MSE	415.32442	R-Square	0.8109
Dependent Mean	1750.98330	Adj R-Sq	0.8104
Coeff Var	23.71950		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-1534.61016	83.57246	-18.36	<.0001
speed_ground	speed_ground	1	41.73170	1.02706	40.63	<.0001

Fig 3: Speed_ground vs distance(Boeing)

Root MSE	307.26984	R-Square	0.8501
Dependent Mean	1323.31696	Adj R-Sq	0.8495
Coeff Var	23.21967		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-2522.89061	85.19508	-29.61	<.0001
speed_ground	speed_ground	1	42.55420	0.86152	49.39	<.0001
height	height	1	14.09773	1.48228	9.51	<.0001

Fig 4: Speed_ground, height vs distance(Airbus)

N	Mean	Std Dev	Std Err	Minimum	Maximum
917	1.25E-12	400.7	13.2311	-816.5	1688.5

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1.25E-12	-25.9668	25.9668	400.7	383.1	419.9

DF	t Value	Pr > |t|
916	0.00	1.0000

Fig 5: T-Test on Residuals (Speed_ground, pitch and height)

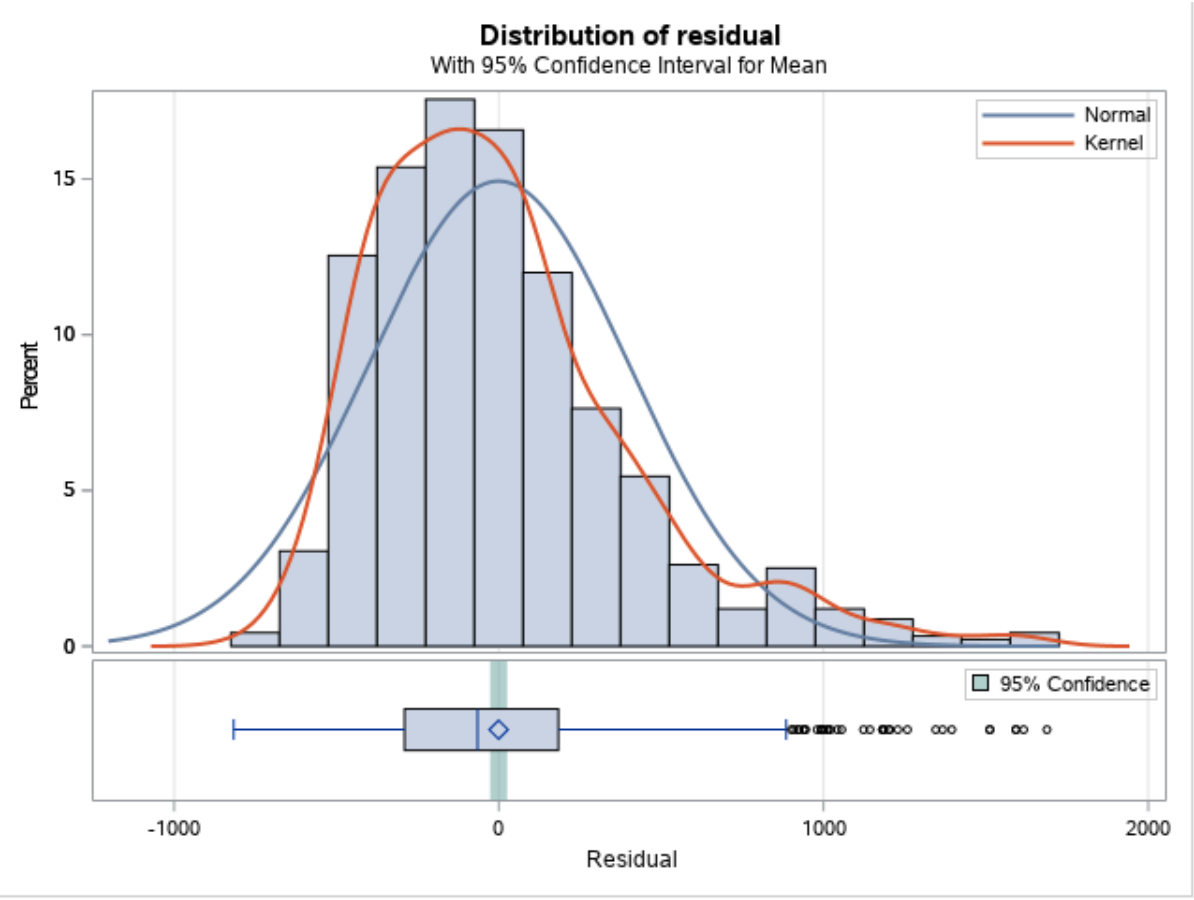


Fig 6: Plot of distribution of residuals

N	Mean	Std Dev	Std Err	Minimum	Maximum
387	2.56E-12	414.8	21.0848	-750.0	1522.1

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2.56E-12	-41.4554	41.4554	414.8	387.5	446.3

DF	t Value	Pr > |t|
386	0.00	1.0000

Fig 7: T-Test on residuals for Boeing aircraft (speed_ground)

N	Mean	Std Dev	Std Err	Minimum	Maximum
444	1.14E-12	306.6	14.5494	-589.2	1450.4

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1.14E-12	-28.5945	28.5945	306.6	287.7	328.2

DF	t Value	Pr > |t|
443	0.00	1.0000

Fig 8: T-Test on residuals for Airbus aircraft (speed_ground, height)

The equations:

Landing distance = -3050+(43.92*speed_ground)+(13.44*height+(170*pitch)

Distance(Boeing) = -1534+(41.7*speed_ground)

Distance(Airbus) = -2522+ (42.55*speed_ground)+(14.1*height)

OBSERVATIONS:

1. From Fig 1 it is seen that though speed_ground had a positive correlation with landing distance individually, it has changed to negative when we add an internally correlated variable speed_air. Therefore, it is better to avoid this combination.
2. The best R-square value is obtained when we have three variables speed_-ground, height and pitch. The p value for all the variables in the LRA is very small and therefore is the best. The residuals are tested using T-Test and the mean was found to be close to 0.
3. When LRA is performed for the Boeing aircraft with speed_ground. The p val-ue is small and hence a very good fit. The residuals are tested for mean 0 and normality. The mean is concluded to be zero after T-Test.
4. LRA was also performed for Airbus aircraft with speed_ground and height as the variables. The p-value for the test was found to be small and the T-Test of the residuals was positive and the mean was concluded to be zero.
5. The aircrafts Boeing and Airbus have a different effect on the landing dis-tance and the type of aircraft does play a role.