Analytics

Understanding Data

ID	Identifier
PAXCOUNT	Number of customer travelling
SALESCHANNEL	Sales channel booking was made on
TRIPTYPEDESC	Trip Type (Round Trip, One Way, Circle Trip)
PURCHASELEAD	Number of days between travel date and booking date
LENGTHOFSTAY	Number of days spent at destination
flight_hour	Hour of day of Flight departure
flight_day	Day of week of Flight departure
ROUTE	Origin-Destination flight route
geoNetwork_country	Country from where booking was made

Understanding Data

BAGGAGE_CATEGORY	Has bought extra baggage in booking
SEAT_CATEGORY	Has bought preferred seat in the booking
FNB_CATEGORY	Has bought in-flight meals
INS_FLAG	Has bought insurance? (Target Variable)
flightDuration_hour	Total duration of flight (in hours)

Understanding 'Passenger Travelled'

For total of 50K booking IDs, there are total of 79,562 passengers. (Avg ~ 1.59)

Trip Category (Avg passengers per booking ID)

- For Round trip category: Avg passengers per booking ID ~ 1.59
- For One Way trip category: Avg passengers per booking ID ~ 1.56

Country-wise (Avg passengers per booking ID)

- For Australia: ~ 1.56
- For Malasiya: ~ 1.44

Understanding 'Sales Channels'

For total of 50K booking IDs, 'Internet' to 'Mobile' booking ~ 89% : ~ 11%

Country-wise (I: M ratio)

- For Australia: ~ 92:8
- For Malasiya: ~ 89 : 11

Majority of bookings has been through 'Internet' sales channels even in Australia and Malaysia.

Please see distribution of sales channels for each country in the data (Tableau file)

Understanding 'Sales Channels'



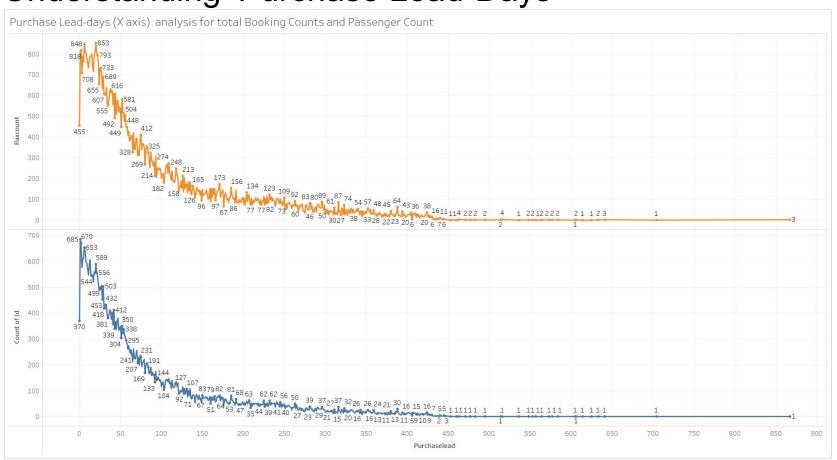
Understanding 'Passengers Travelled'

For total of 50K booking IDs, total 79,562 passengers travelled (i.e., 1.6 passenger per booking).

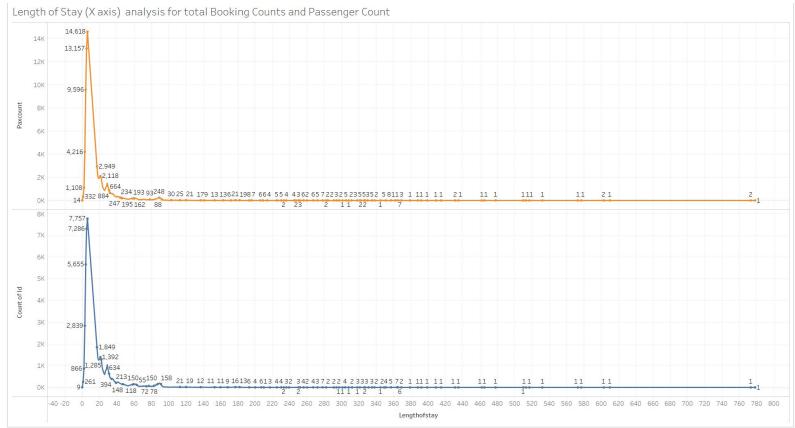


Understanding 'Purchase Lead-Days'

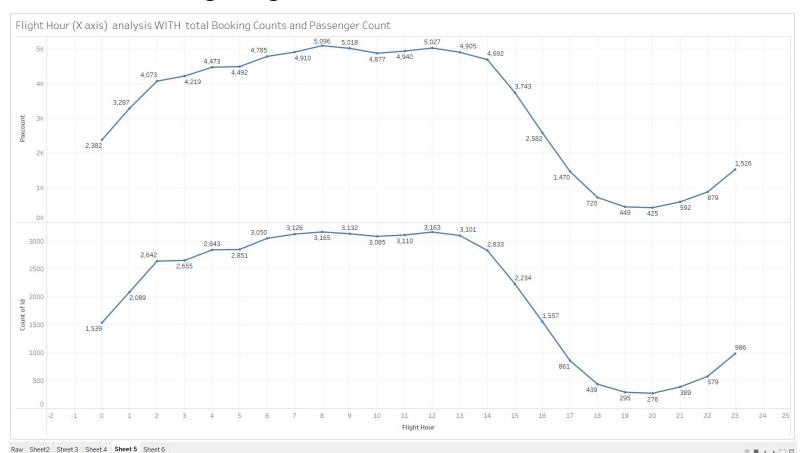
Raw Sheet2 Sheet3 Sheet4 Sheet5 Sheet6 Sheet7



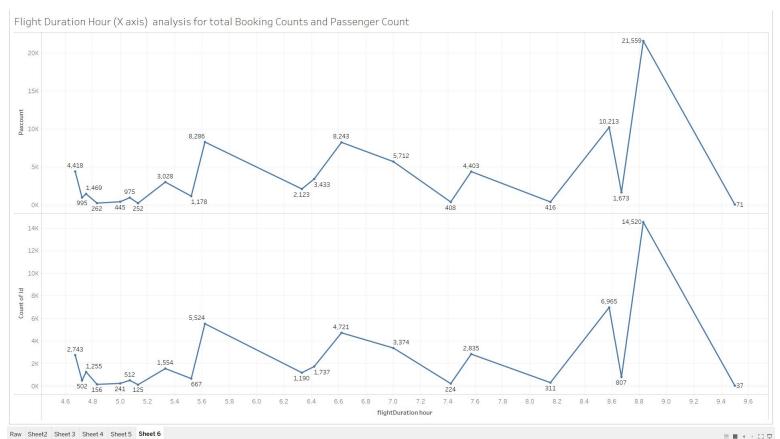
Understanding 'Length of Stay'



Understanding Flight Hours



Understanding flightDuration Hours



Extra Baggage Claims

- Out of 50K bookings, about 67% claimed extra baggages in booking (i.e., 33,439 bookings).
- 90% (i.e., 30,085) of these 'extra baggage claims' are from sale's channel = Internet.
- 99% (i.e., 33,075) of these 'extra baggage claims' are from Trip Type = Round Trip.
- Higher the length of stay or higher the number of the passengers -> more claims for extra baggage.
 - Avg # of passengers per bookings are 1.7 for 'extra baggage claim'.
 - Avg # length of stay for 'extra baggage claims' were about 27 days (much higher than 14.5 days, in case of no extra baggage claim).

Ancillary Services:

- Out of 33,439 extra baggage claims, only ~36% choose 'preferred' seats.
- Out of 33,439 extra baggage claims, only ~50% bought FnB in the flight.
- Out of 33,439 extra baggage claims, only ~18% bought 'insurance' during the flight.
- Only 1,631 booking had bought all the four ancillary services (extra baggage claim, preferred seats, FnB, insurance).

Case: when no extra baggage claim

df0.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	16561	2	Internet	14297
TRIPTYPEDESC	16561	3	RoundTrip	16424
flight_day	16561	7	Mon	2730
ROUTE	16561	706	AKLKUL	867
geoNetwork_country	16561	82	Australia	5343

Case: when there is an extra baggage claim

dfb.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	33439	2	Internet	30085
TRIPTYPEDESC	33439	3	RoundTrip	33073
flight_day	33439	7	Mon	5372
ROUTE	33439	764	AKLKUL	1813
geoNetwork_country	33439	90	Australia	12529

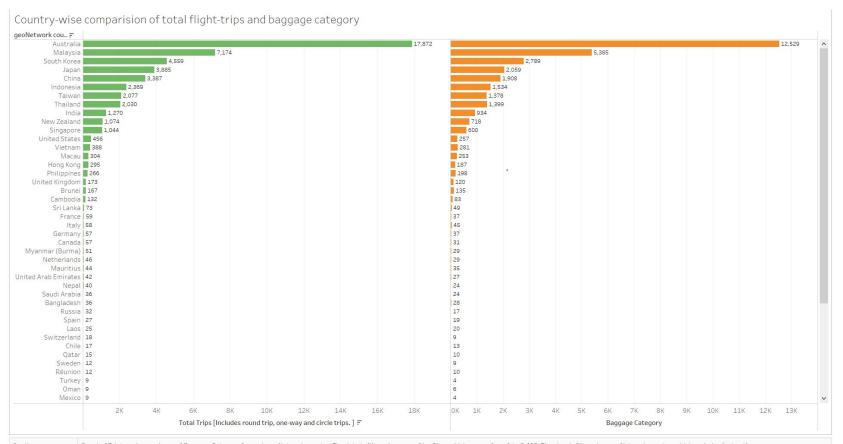
Case: when no extra baggage claim

Case: when	there is a	an extra l	baggage cl	aim
------------	------------	------------	------------	-----

	count	mean	std	min	25%	50%	75%	max		count	mean	std	min	25%	50%	75%	max
ld	16561.0	28499.4	13708.6	2.0	18358.0	30428.0	39412.0	50000.0	ld	33439.0	23267.6	14470.2	1.0	11156.5	21365.0	35989.0	49999.0
PAXCOUNT	16561.0	1.4	0.8	1.0	1.0	1.0	2.0	9.0	PAXCOUNT	33439.0	1.7	1.1	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	16561.0	87.8	97.9	0.0	19.0	49.0	119.0	867.0	PURCHASELEAD	33439.0	83.5	86.5	0.0	22.0	52.0	113.0	641.0
LENGTHOFSTAY	16561.0	14.5	25.5	0.0	4.0	6.0	20.0	778.0	LENGTHOFSTAY	33439.0	27.3	36.6	0.0	5.0	19.0	31.0	773.0
flight_hour	16561.0	9.2	5.5	0.0	5.0	9.0	13.0	23.0	flight_hour	33439.0	9.0	5.4	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	16561.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	BAGGAGE_CATEGORY	33439.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0
SEAT_CATEGORY	16561.0	0.2	0.4	0.0	0.0	0.0	0.0	1.0	SEAT_CATEGORY	33439.0	0.4	0.5	0.0	0.0	0.0	1.0	1.0
FNB_CATEGORY	16561.0	0.3	0.4	0.0	0.0	0.0	1.0	1.0	FNB_CATEGORY	33439.0	0.5	0.5	0.0	0.0	1.0	1.0	1.0
INS_FLAG	16561.0	0.1	0.3	0.0	0.0	0.0	0.0	1.0	INS_FLAG	33439.0	0.2	0.4	0.0	0.0	0.0	0.0	1.0
flightDuration_hour	16561.0	7.2	1.5	4.7	5.6	7.0	8.7	9.5	flightDuration_hour	33439.0	7.3	1.5	4.7	5.6	7.6	8.8	9.5

- More the passengers per booking -> more the need of extra baggage claim
- More the length of stay -> more the need of the extra baggage claim.

Country ->	AUSTF	RALIA	MALAYSIA			
Features	BAGGAGE_CATEGORY = 0	BAGGAGE_CATEGORY = 1	BAGGAGE_CATEGORY = 0	BAGGAGE_CATEGORY = 1		
PAXCOUNT	1.3	1.5	1.4	1.8		
PURCHASELEAD	73.5	88.7	96	86		
LENGTHOFSTAY	20	33	16	27		
flightDuration_hour	7.9	8	7	6.8		



Preferred Seats Purchase

- Out of 50K bookings, about 30% purchased preferred seats in booking (i.e., 14,848 bookings).
- 88% of these bookings are from sale's channel = Internet.
- 99% of these bookings are from Trip Type = Round Trip.
- Less dependent on the length of stay or the number of the passengers.

Ancillary Services:

- Out of 14,848 such bookings, only ~82% brought extra baggages.
- Out of 14,848 such bookings, only ~67% bought FnB in the flight.
- Out of 14,848 such bookings, only ~11% bought 'insurance' during the flight.
- Only 1,631 booking had bought all the four ancillary services (extra baggage claim, preferred seats, FnB, insurance).

Case: preferred seats

dfs.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	14848	2	Internet	12942
TRIPTYPEDESC	14848	3	RoundTrip	14715
flight_day	14848	7	Mon	2464
ROUTE	14848	657	AKLKUL	611
geoNetwork_country	14848	78	Australia	6150

Case: no preferred seats

dfs0.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	35152	2	Internet	31440
TRIPTYPEDESC	35152	3	RoundTrip	34782
flight_day	35152	7	Mon	5638
ROUTE	35152	779	AKLKUL	2069
geoNetwork_country	35152	94	Australia	11722

Case: preferred seats

	und(1)							
	count	mean	std	min	25%	50%	75%	max
ld	14848.0	24577.2	14301.0	3.0	12249.2	24004.0	37005.8	50000.0
PAXCOUNT	14848.0	1.6	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	14848.0	84.3	87.1	0.0	22.0	52.0	116.0	584.0
LENGTHOFSTAY	14848.0	24.7	37.5	0.0	5.0	17.0	29.0	610.0
flight_hour	14848.0	9.2	5.5	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	14848.0	8.0	0.4	0.0	1.0	1.0	1.0	1.0
SEAT_CATEGORY	14848.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0
FNB_CATEGORY	14848.0	0.7	0.5	0.0	0.0	1.0	1.0	1.0
INS_FLAG	14848.0	0.2	0.4	0.0	0.0	0.0	0.0	1.0
flightDuration_hour	14848.0	7.5	1.4	4.7	6.4	8.6	8.8	9.5

Case: no preferred seats

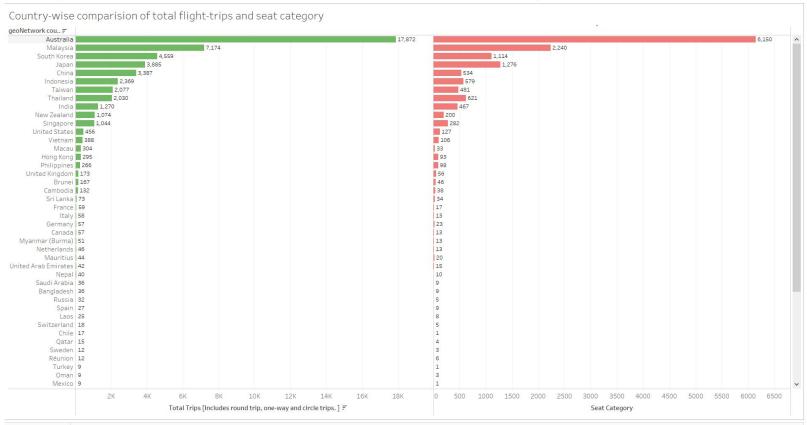
dfs0.describe().T.round(1)

	count	mean	std	min	25%	50%	75%	max
ld	35152.0	25179.3	14486.2	1.0	12666.8	25440.5	37689.2	49999.0
PAXCOUNT	35152.0	1.6	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	35152.0	85.2	91.8	0.0	21.0	51.0	115.0	867.0
LENGTHOFSTAY	35152.0	22.3	32.2	0.0	5.0	6.0	28.0	778.0
flight_hour	35152.0	9.0	5.4	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	35152.0	0.6	0.5	0.0	0.0	1.0	1.0	1.0
SEAT_CATEGORY	35152.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FNB_CATEGORY	35152.0	0.3	0.5	0.0	0.0	0.0	1.0	1.0
INS_FLAG	35152.0	0.1	0.3	0.0	0.0	0.0	0.0	1.0
flightDuration_hour	35152.0	7.2	1.5	4.7	5.6	7.0	8.8	9.5

- More the passengers per booking -> higher chances of purchase of preferred seats
- More the purchase lead-days -> higher chances of purchase of preferred seats

Country ->	AUSTF	RALIA	MALAYSIA			
Features	SEAT_CATEGORY = 0	SEAT_CATEGORY = 1	SEAT_CATEGORY = 0	SEAT_CATEGORY = 1		
PAXCOUNT	1.4	1.5	1.6	1.8		
PURCHASELEAD	76	82	85	92		
LENGTHOFSTAY	28	30	26	20		
flightDuration_hour	7.9	8.0	6.8	7.2		

Caption



Count of Triptypedesc and sum of Seat Category for each geoNetwork country, The data is filtered on sum of Ins Flag, which ranges from 1 to 2,468. The view is filtered on geoNetwork country, which excludes (not set).

In-flight FnB Purchase

- Out of 50K bookings, about ~43% purchased preferred seats in booking (i.e., 21,357 bookings).
- ~90% of these bookings are from sale's channel = Internet.
- ~99% of these bookings are from Trip Type = Round Trip.

Ancillary Services:

- Out of 21,357 such purchases, only ~80% brought extra baggages.
- Out of 21,357 such purchases, only ~47% bought preferred seats.
- Out of 21,357 such purchases, only ~16% bought 'insurance' during the flight.
- Only 1,631 booking had bought all the four ancillary services (extra baggage claim, preferred seats, FnB, insurance).

Case: purchased FnB

dfFNB.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	21357	2	Internet	19131
TRIPTYPEDESC	21357	3	RoundTrip	21177
flight_day	21357	7	Mon	3423
ROUTE	21357	711	AKLKUL	1203
geoNetwork_country	21357	83	Australia	9068

Case: not purchased FnB

dfFNB0.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	28643	2	Internet	25251
TRIPTYPEDESC	28643	3	RoundTrip	28320
flight_day	28643	7	Mon	4679
ROUTE	28643	760	AKLKUL	1477
geoNetwork_country	28643	93	Australia	8804

Case: purchased FnB

dfFNB.describe().T.round()

	count	mean	std	min	25%	50%	75%	max
ld	21357.0	23224.0	14492.0	4.0	10401.0	21802.0	35634.0	49999.0
PAXCOUNT	21357.0	2.0	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	21357.0	83.0	86.0	0.0	22.0	52.0	112.0	867.0
LENGTHOFSTAY	21357.0	27.0	38.0	0.0	5.0	19.0	31.0	773.0
flight_hour	21357.0	9.0	5.0	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	21357.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0
SEAT_CATEGORY	21357.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
FNB_CATEGORY	21357.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0
INS_FLAG	21357.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
flightDuration_hour	21357.0	8.0	1.0	5.0	6.0	9.0	9.0	10.0

Case: not purchased FnB

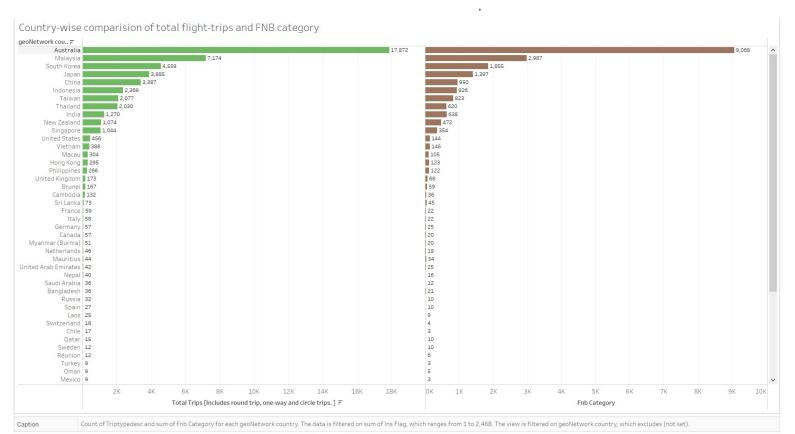
dfFNB0.describe().T.round()

	count	mean	std	min	25%	50%	75%	max
ld	28643.0	26325.0	14247.0	1.0	14228.0	27421.0	38422.0	50000.0
PAXCOUNT	28643.0	2.0	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	28643.0	87.0	94.0	0.0	21.0	51.0	118.0	704.0
LENGTHOFSTAY	28643.0	20.0	30.0	0.0	5.0	6.0	26.0	778.0
flight_hour	28643.0	9.0	5.0	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	28643.0	1.0	0.0	0.0	0.0	1.0	1.0	1.0
SEAT_CATEGORY	28643.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
FNB_CATEGORY	28643.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
INS_FLAG	28643.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
flightDuration_hour	28643.0	7.0	2.0	5.0	6.0	7.0	9.0	10.0

Varied trends w.r.t two countries.

Country ->	AUSTF	RALIA	MALAYSIA		
Features	FNB_CATEGORY = 0	FNB_CATEGORY = 1	FNB_CATEGORY = 0	FNB_CATEGORY = 1	
PAXCOUNT	1.4	1.5	1.6	1.8	
PURCHASELEAD	80	78	86	92.5	
LENGTHOFSTAY	26	32	25	24	
flightDuration_hour	7.8	8.1	6.7	7.2	

Raw Sheet2 Sheet3 Sheet4 Sheet5 Sheet6 Sheet7 Sheet8 Sheet9 Sheet10 Sheet11 Sheet 12



悪量する日草

Insurance Purchase

- Out of 50K bookings, about ~15% purchased insurance (i.e., 7,478 bookings).
- ~90% of these bookings are from sale's channel = Internet.
- ~99% of these bookings are from Trip Type = Round Trip.

Ancillary Services:

- Out of 7,478 such purchases, only ~75% brought extra baggages.
- Out of 7,478 such purchases, only ~36% bought preferred seats.
- Out of 7,478 such purchases, only ~46% bought 'FnB.
- Only 1,631 booking had bought all the four ancillary services (extra baggage claim, preferred seats, FnB, insurance).

Case: Insurance purchased

dfINS.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	7478	2	Internet	6869
TRIPTYPEDESC	7478	3	RoundTrip	7453
flight_day	7478	7	Wed	1252
ROUTE	7478	549	AKLKUL	567
geoNetwork_country	7478	58	Malaysia	2468

Case: No insurance purchased

dfINS0.describe(include = [object]).T.round()

	count	unique	top	freq
SALESCHANNEL	42522	2	Internet	37513
TRIPTYPEDESC	42522	3	RoundTrip	42044
flight_day	42522	7	Mon	6898
ROUTE	42522	786	AKLKUL	2113
geoNetwork_country	42522	100	Australia	16972

Case: Insurance purchased

dfINS.describe().T.round()

	count	mean	std	min	25%	50%	75%	max
ld	7478.0	28580.0	14449.0	83.0	17338.0	30596.0	41076.0	49988.0
PAXCOUNT	7478.0	2.0	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	7478.0	80.0	90.0	0.0	18.0	46.0	106.0	867.0
LENGTHOFSTAY	7478.0	20.0	34.0	1.0	5.0	6.0	23.0	778.0
flight_hour	7478.0	9.0	5.0	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	7478.0	1.0	0.0	0.0	0.0	1.0	1.0	1.0
SEAT_CATEGORY	7478.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
FNB_CATEGORY	7478.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
INS_FLAG	7478.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0
flightDuration_hour	7478.0	7.0	2.0	5.0	6.0	7.0	9.0	10.0

Case: No insurance purchased

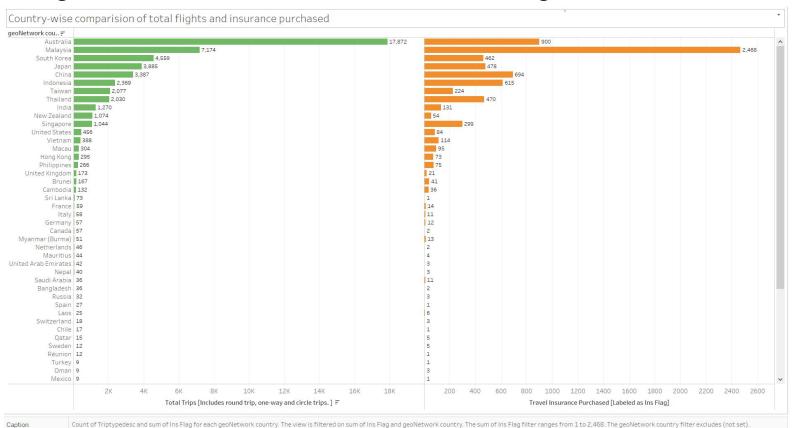
dfINS0.describe().T.round()

	count	mean	std	min	25%	50%	75%	max
ld	42522.0	24371.0	14339.0	1.0	11873.0	23900.0	36769.0	50000.0
PAXCOUNT	42522.0	2.0	1.0	1.0	1.0	1.0	2.0	9.0
PURCHASELEAD	42522.0	86.0	91.0	0.0	22.0	52.0	117.0	704.0
LENGTHOFSTAY	42522.0	24.0	34.0	0.0	5.0	17.0	29.0	773.0
flight_hour	42522.0	9.0	5.0	0.0	5.0	9.0	13.0	23.0
BAGGAGE_CATEGORY	42522.0	1.0	0.0	0.0	0.0	1.0	1.0	1.0
SEAT_CATEGORY	42522.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
FNB_CATEGORY	42522.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
INS_FLAG	42522.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
flightDuration_hour	42522.0	7.0	1.0	5.0	6.0	8.0	9.0	10.0

Varied trends w.r.t two countries.

Country ->	AUSTF	RALIA	MALAYSIA		
Features	INS_FLAG = 0	INS_FLAG = 1	INS_FLAG = 0	INS_FLAG = 1	
PAXCOUNT	1.5	1.3	1.7	1.6	
PURCHASELEAD	80	60	88	87	
LENGTHOFSTAY	29	35	27	20	
flightDuration_hour	8	8	6.9	6.8	

Raw Sheet2 Sheet3 Sheet4 Sheet5 Sheet6 Sheet7 Sheet8 Sheet9 Sheet10 Sheet11 Sheet12



悪量・・口草