

Universidad del Valle de Guatemala

Bases de Datos

Ing. Héctor Hurtarte

Sección: 10



## ***Proyecto 1***

# ***Consulta y análisis de Base de Datos de Aviación del Bureau of National Statistics***

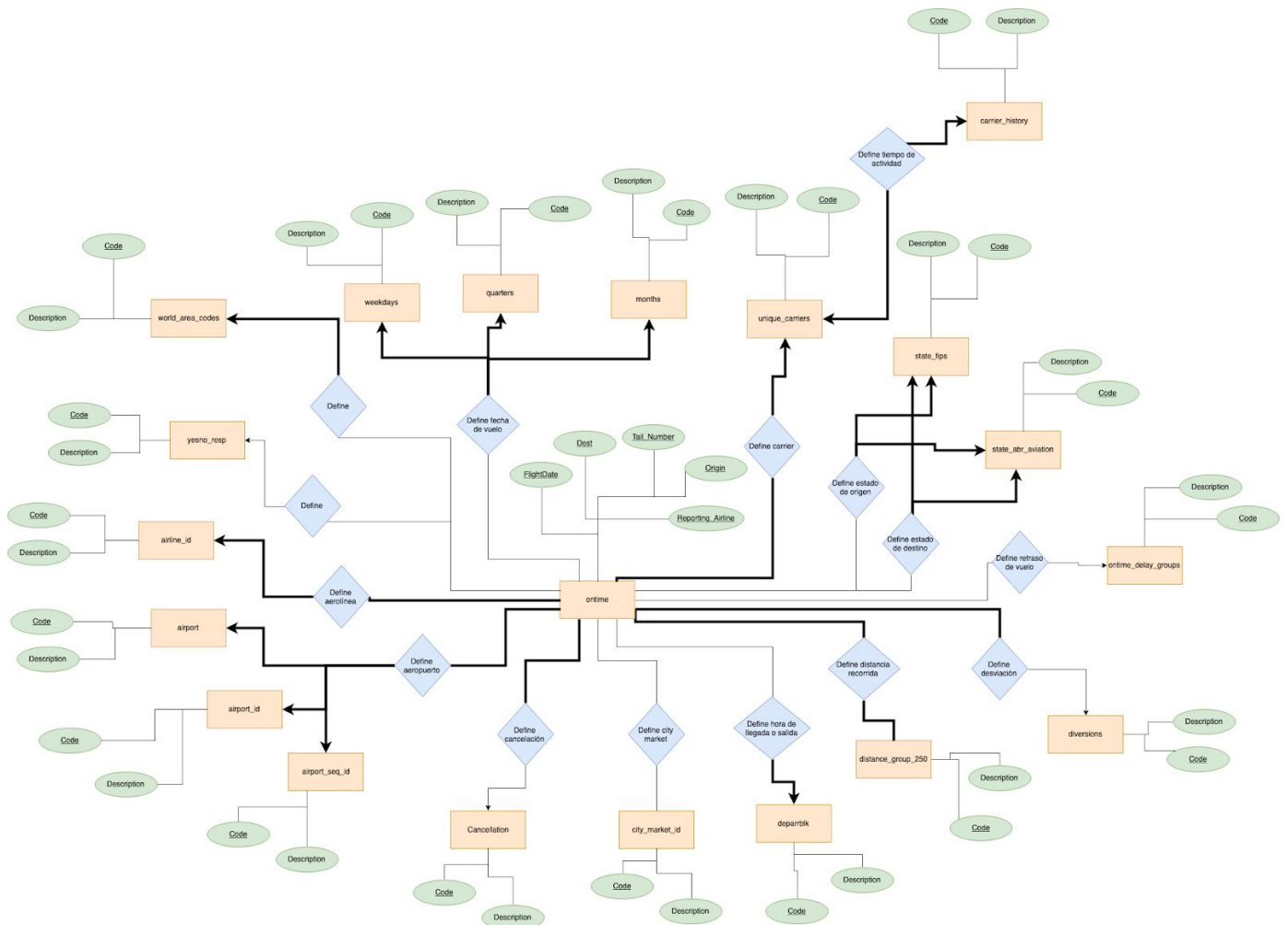
María Fernanda López - 17160

Ana Lucía Hernández - 17138

Guatemala, 27 de febrero de 2019

## I. Modelo de Datos

### A. Diagrama Entidad-Relación



Si se desea apreciar el diagrama Entidad-Relación más de cerca, utilizar el archivo DiagramaEntidadRelacion.svg

### B. Diccionario de la entidad “ontime”

Atributos que conforman la llave primaria:

- |                      |               |         |
|----------------------|---------------|---------|
| 1. Reporting_Airline | 3. FlightDate | 5. Dest |
| 2. Tail_Number       | 4. Origin     |         |

Demás atributos de la entidad:

- |         |            |          |
|---------|------------|----------|
| 1. Year | 2. Quarter | 3. Month |
|---------|------------|----------|

4. DayOfMonth	33. WheelsOn	64. Div1Airport
5. DayOfWeek	34. TaxiIn	65. Div1AirportId
6. DOT_ID_Reporting_Airline	35. CrsArrTime	66. Div1AirportSeqId
7. IATA_CODE_Report ing_Airline	36. ArrTime	67. Div1WheelsOn
8. Flight_Number_Repo rting_Airline	37. ArrDelay	68. Div1TotalGtime
9. OriginAirportId	38. ArrDelayMinutes	69. Div1WheelsOff
10. OriginCityMarketId	39. ArrDel15	70. Div1TailNum
11. OriginCityName	40. ArrDelayGroup	71. Div2Airport
12. OriginStateAbr	41. ArrTimeBlk	72. Div2AirportId
13. OriginStateFips	42. Cancelled	73. Div2AirportSeqId
14. OriginStateNm	43. CancellationCode	74. Div2WheelsOn
15. OriginWac	44. Diverted	75. Div2TotalGtime
16. DestAirportId	45. CrsElapsedTime	76. Div2WheelsOff
17. DestAirportSeqId	46. ActualElapsedTime	77. Div2TailNum
18. DestCityMarketId	47. AirTime	78. Div3Airport
19. DestCityName	48. Flights	79. Div3AirportId
20. DestStateAbr	49. Distance	80. Div3AirportSeqId
21. DestStateFips	50. DistanceGroup	81. Div3WheelsOn
22. DestStateNm	51. CarrierDelay	82. Div3TotalGtime
23. DestWac	52. WeatherDelay	83. Div3WheelsOff
24. CrsDepTime	53. NasDelay	84. Div3TailNum
25. DepTime	54. SecurityDelay	85. Div4Airport
26. DepDelay	55. LateAircraftDelay	86. Div4AirportId
27. DepDelayMinutes	56. FirstDepTime	87. Div4AirportSeqId
28. DepDelay15	57. TotalAddGTime	88. Div4WheelsOn
29. DepDelayGroup	58. LongestAddGTime	89. Div4TotalGtime
30. DepTimeBlk	59. DivAirportLandings	90. Div4WheelsOff
31. TaxiOut	60. DivReachedDest	91. Div4TailNum
32. WheelsOff	61. DivActualElapsedTi me	92. Div5Airport
	62. DivArrDelay	93. Div5AirportId
	63. DivDistance	94. Div5AirportSeqId
		95. Div5WheelsOn

96. Div5TotalGtime

97. Div5WheelsOff

98. Div5TailNum

## II. Análisis de Base de Datos

### A. Planteamiento de incógnitas

1. ¿Cuáles son las cinco aerolíneas que han cancelado - proporcionalmente - menos vuelos?
2. ¿Cuáles son las cinco aerolíneas que tienen la mayor cantidad de estados como destino?
3. ¿Cuáles son las aerolíneas que mayor tiempo de adelanto han logrado tener en sus vuelos?
4. ¿Qué 5 aerolíneas tienen la menor aportación porcentual al total de tiempos de retraso de llegada de todos los vuelos?
5. ¿Cuál es la frecuencia de vuelos para las aerolíneas que tienen menor tiempo de retraso y menor cantidad de vuelos cancelados?
6. ¿Cuál es la proporción o razón de distancia total recorrida (de todas las aeronaves) respecto a la cantidad de vuelos?
7. ¿Cuáles 5 aerolíneas tienen predichas el mayor crecimiento en operaciones?
8. ¿Cuáles aerolíneas tienen la mayor cantidad de vuelos en los estados más poblados?
9. De las aerolíneas que tienen menor tiempo de retraso y llegan antes, ¿Cuál es su velocidad promedio en sus vuelos?
10. Para las aerolíneas con los vuelos que recorrieron las distancias mayores, ¿Cuál es su eficiencia (tiempo en el aire) respecto a la distancia que recorren?

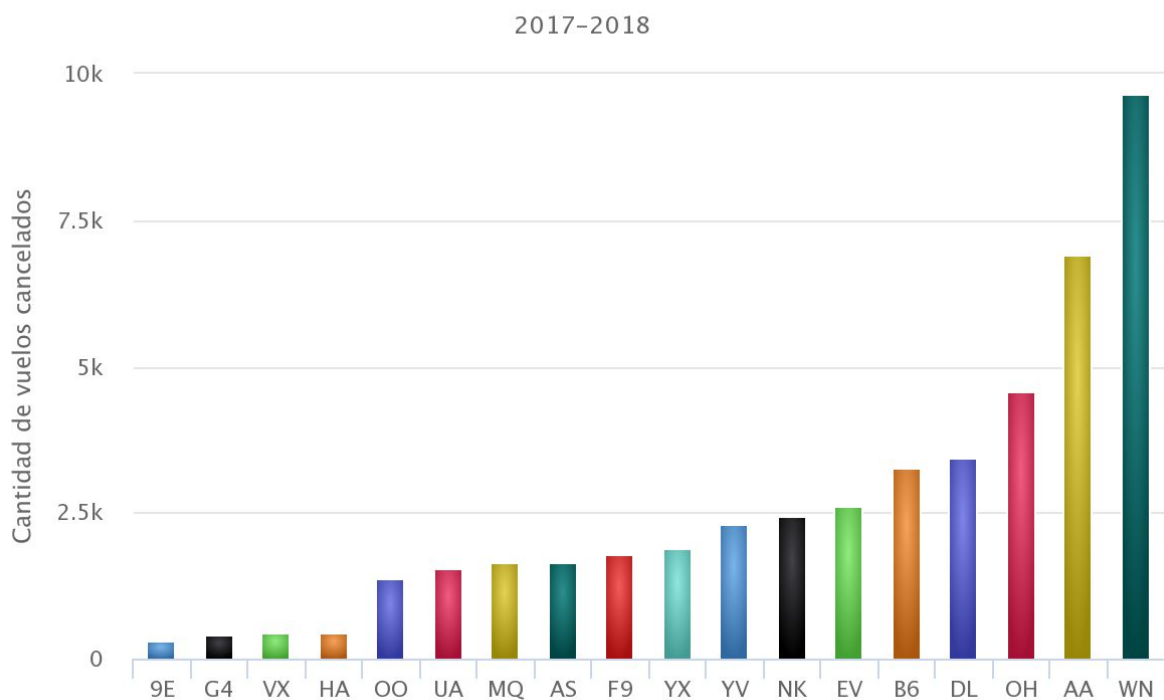
### B. Resultados obtenidos con los queries

#### 1. Primer Query

```
SELECT Reporting_Airline, (COUNT(*)*1.0*100)/Cancelados as Porcentaje
FROM ontime,
(
    SELECT COUNT(*) as Cancelados
    FROM ontime
    WHERE Cancelled = 1 and CancellationCode = "A"
)
WHERE Cancelled = 1 and CancellationCode = "A"
GROUP BY Reporting_Airline
ORDER BY Porcentaje ASC
```

Reporting_Airline	Porcentaje
9E	0.681102615
G4	0.852449185
VX	0.927413309
HA	0.970249952
OO	2.972862987
UA	3.277003148
MQ	3.504037354
AS	3.529739339
F9	3.840304997
YX	4.033069888
YV	4.954057701
NK	5.247488702
EV	5.620167491
B6	6.986656386
DL	7.340058686
OH	9.768896314
AA	14.81933646
WN	20.67510549

## Aerolíneas con menor cantidad de vuelos cancelos



## 2. Segundo Query

```

SELECT unique_carriers.Description as aerolinea, COUNT(DISTINCT DestStateNm) as
cantidad_estados
FROM ontime
INNER JOIN unique_carriers ON IATA_CODE_Reporting_Airline = code
GROUP BY IATA_CODE_Reporting_Airline
ORDER BY COUNT(DISTINCT DestStateNm) DESC

```

Aerolínea	CantidadEstados
Delta Air Lines Inc.	51
SkyWest Airlines Inc.	48
American Airlines Inc.	46
United Air Lines Inc.	46
Allegiant Air	44
Frontier Airlines Inc.	42
ExpressJet Airlines Inc.	41
Southwest Airlines Co.	41
Mesa Airlines Inc.	41
Envoy Air	39
Republic Airline	39
Endeavor Air Inc.	37
PSA Airlines Inc.	33
JetBlue Airways	32
Alaska Airlines Inc.	31
Spirit Air Lines	27
Virgin America	20
Hawaiian Airlines Inc.	8

### 3. Tercer Query

- a. Para obtener las aerolíneas con el mayor tiempo de adelanto acumulado

```

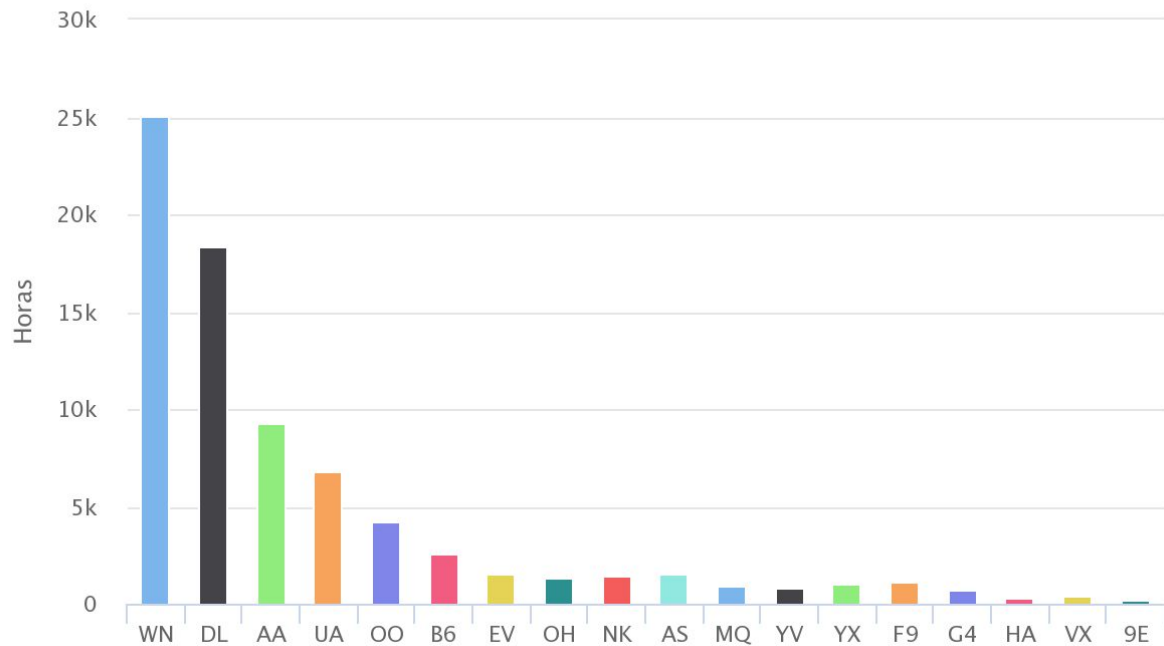
SELECT DISTINCT unique_carriers.Description as aerolinea, abs(SUM(ArrDelay)) as
TiempoGanado
FROM ontime
JOIN unique_carriers ON unique_carriers.Code = ontime.Reporting_Airline
WHERE DepDelay = 0 and ArrDelay < 0
GROUP BY Reporting_Airline
ORDER BY TiempoGanado DESC

```

Aerolínea	TiempoGanado
Southwest Airlines Co.	1501573
Delta Air Lines Inc.	1102235
American Airlines Inc.	554220
United Air Lines Inc.	410217
SkyWest Airlines Inc.	251206
JetBlue Airways	154461
Alaska Airlines Inc.	95371
ExpressJet Airlines Inc.	90527
Spirit Air Lines	85496
PSA Airlines Inc.	78411
Frontier Airlines Inc.	65849
Republic Airline	62388
Envoy Air	54984
Mesa Airlines Inc.	49184
Allegiant Air	42324
Virgin America	23592
Hawaiian Airlines Inc.	19221
Endeavor Air Inc.	10958

## Total en horas de tiempo de adelanto

2017-2018

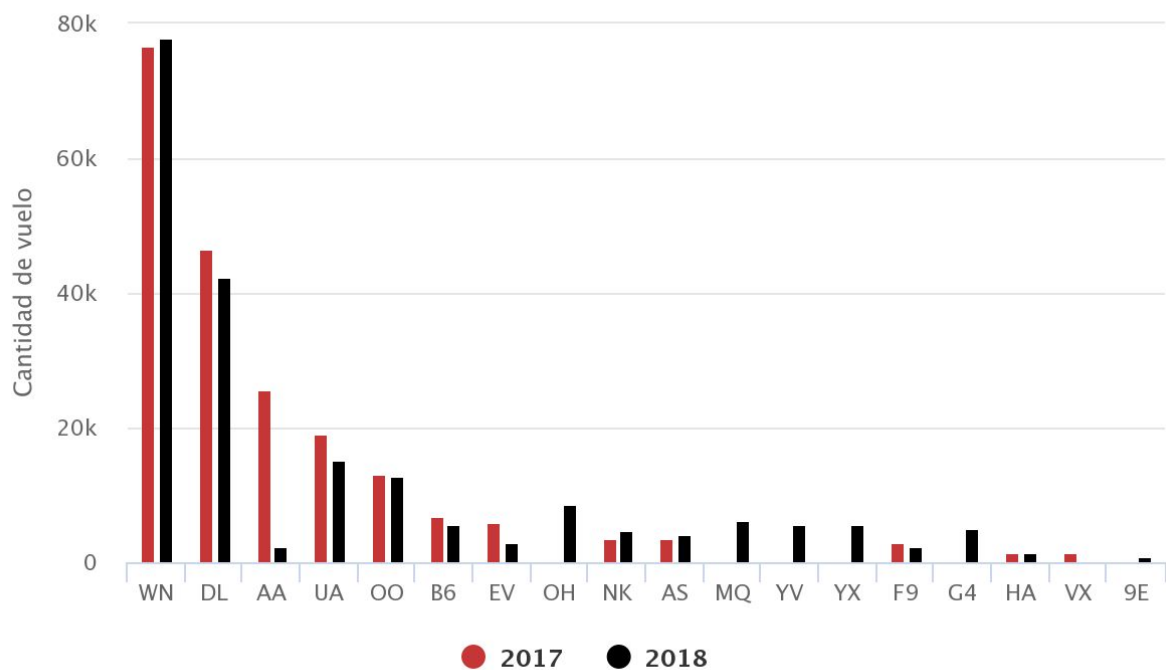


- b. Para obtener la cantidad de vuelos por aerolínea que lograron salir a tiempo y llegar con tiempo adelantado (antes de lo previsto)

```
SELECT DISTINCT unique_carriers.Description as aerolinea, SUM(CASE WHEN
"Year"="2017" THEN 1 ELSE 0 END) as "CantidadVuelos2017", SUM(CASE WHEN
"Year"="2018" THEN 1 ELSE 0 END) as "CantidadVuelos2018"
FROM ontime
INNER JOIN unique_carriers ON unique_carriers.Code = ontime.Reporting_Airline
WHERE DepDelay = 0 and ArrDelay < 0
GROUP BY Reporting_Airline
ORDER BY ("2017"+"2018") DESC
```

Aerolínea	CantidadVuelos2017	CantidadVuelos2018
Endeavor Air Inc.	0	998
American Airlines Inc.	25723	23338
Alaska Airlines Inc.	3457	4254
JetBlue Airways	6854	5637
Delta Air Lines Inc.	46464	42483
ExpressJet Airlines Inc.	6004	2943
Frontier Airlines Inc.	2996	2519
Allegiant Air	0	5051
Hawaiian Airlines Inc.	1419	1433
Envoy Air	0	6123
Spirit Air Lines	3462	4642
PSA Airlines Inc.	0	8716
SkyWest Airlines Inc.	13086	12903
United Air Lines Inc.	19252	15319
Virgin America	1529	336
Southwest Airlines Co.	76667	77946
Mesa Airlines Inc.	0	5564
Republic Airline	0	5554

## Vuelos que salen a tiempo y llegan antes



### 4. Cuarto Query

```

SELECT (SUM(CarrierDelay)*1.0*100)/(RetrasosAerolineas) as PorcentajeRetraso,
U.Description
FROM ontime,
(

```



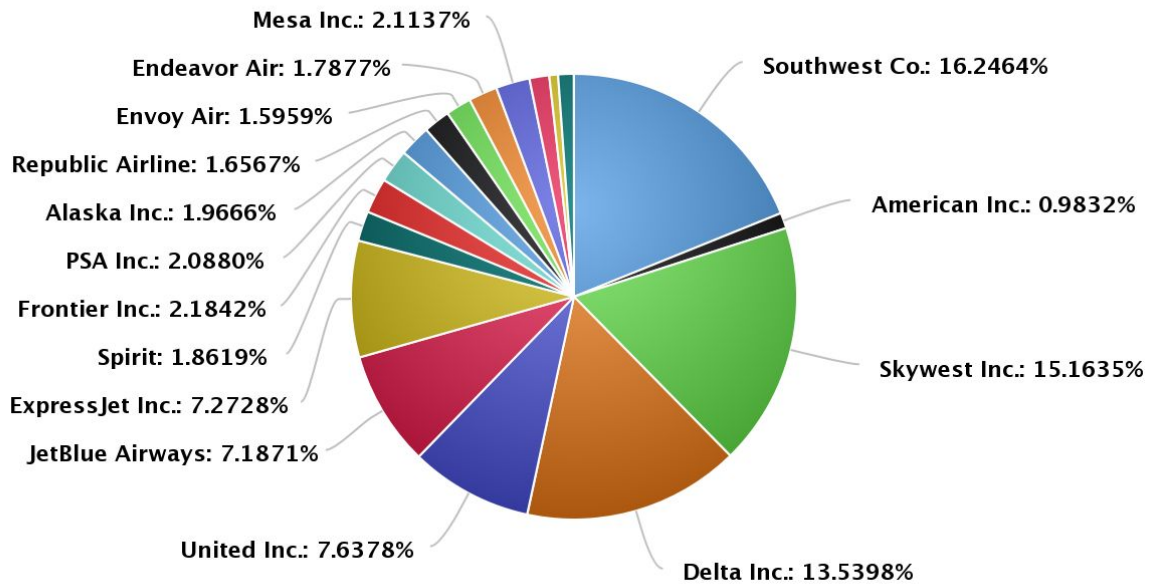
```

SELECT SUM(CarrierDelay) as RetrasosAerolineas
FROM ontime
WHERE CarrierDelay >0
)
JOIN unique_carriers U ON U.code = ontime.Reporting_Airline
WHERE CarrierDelay >0
GROUP BY Reporting_Airline
ORDER BY PorcentajeRetraso ASC

```

Description	PorcentajeRetraso
Virgin America	0.544561815
Hawaiian Airlines Inc.	0.983262556
Allegiant Air	1.227505075
Envoy Air	1.595996964
Republic Airline	1.656732789
Endeavor Air Inc.	1.787774851
Spirit Air Lines	1.861961045
Alaska Airlines Inc.	1.966644632
PSA Airlines Inc.	2.088063161
Mesa Airlines Inc.	2.113708946
Frontier Airlines Inc.	2.184267503
JetBlue Airways	7.187105759
ExpressJet Airlines Inc.	7.272814518
United Air Lines Inc.	7.637877969
Delta Air Lines Inc.	13.53984148
American Airlines Inc.	14.94192692
SkyWest Airlines Inc.	15.16354262
Southwest Airlines Co.	16.2464114

## Aportación porcentual al tiempo de retraso

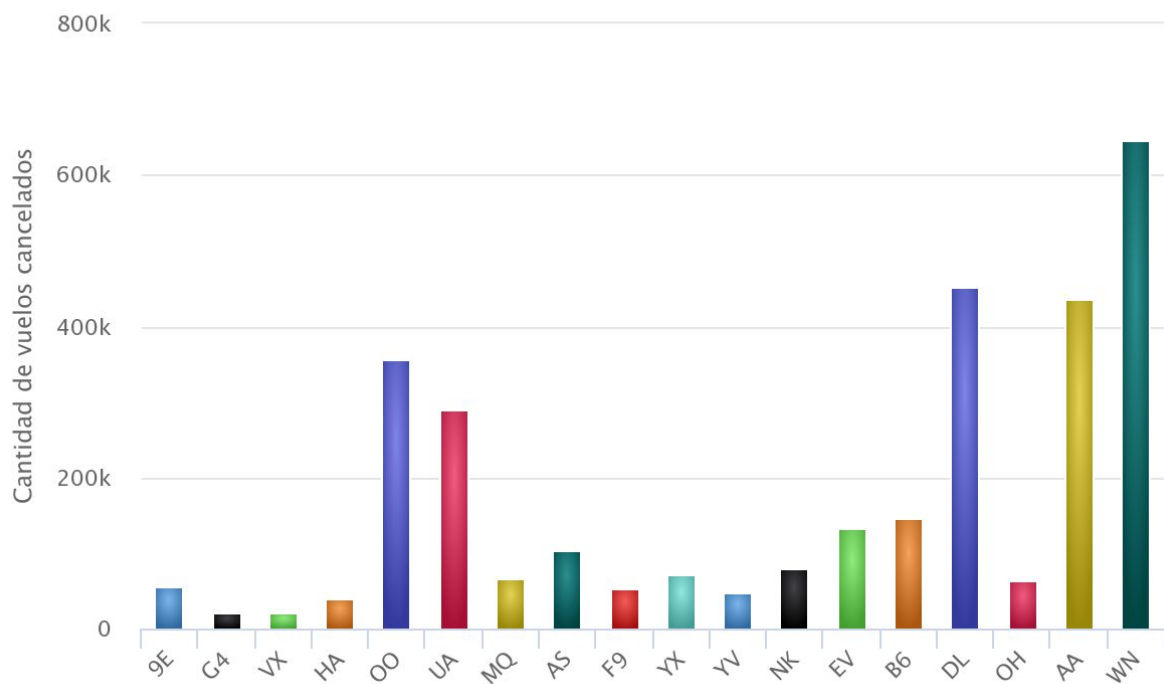


### 5. Quinto Query

```
SELECT
    Reporting_Airline,
    SUM(CASE WHEN Cancelled = 1 AND CancellationCode = 'A' THEN 1 ELSE 0 END) AS
    Cancelado,
    (SUM(CASE WHEN Quarter = 1 THEN 1 ELSE 0 END) +
    SUM(CASE WHEN Quarter = 2 THEN 1 ELSE 0 END) +
    SUM(CASE WHEN Quarter = 3 THEN 1 ELSE 0 END) +
    SUM(CASE WHEN Quarter = 4 THEN 1 ELSE 0 END))/4 as CantVuelosPorCuarto
FROM ontime ontime_externo
GROUP BY Reporting_Airline
ORDER BY Cancelado ASC;
```

Reporting_Airline	Cancelados	CantVuelosPorCuarto
9E	318	56306
G4	398	22137
VX	433	22332
HA	453	39372
OO	1388	356783
UA	1530	290752
MQ	1636	67601
AS	1648	103008
F9	1793	53545
YX	1883	72612
YV	2313	49140
NK	2450	79855
EV	2624	133335
B6	3262	145444
DL	3427	452897
OH	4561	63849
AA	6919	437009
WN	9653	647093

## Aerolíneas con menor cantidad de vuelos por cuarto



### 6. Sexto Query

```

SELECT Reporting_Airline as Aerolinea, (SUM(Distance)/COUNT(*)) as Razon
FROM ontime
GROUP BY Reporting_Airline
ORDER BY Razon ASC

```

Aerolínea	Razon
OH	382.9295173
MQ	433.2349254
EV	465.3439232
OO	473.3308645
9E	475.9989166
YV	550.1730701
YX	594.3277696
HA	644.1620758
WN	747.438323
DL	879.1340568
G4	881.8799648
NK	1018.517601
AA	1021.395659
B6	1053.674029
F9	1059.990055
UA	1185.179944
AS	1277.566661
VX	1482.031972

## 7. Séptimo Query

```

SELECT
    Reporting_Airline,
    (((SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 1 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 1 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
    ((SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
    ((SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
    ((SUM(CASE WHEN Quarter = 1 and "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)))/4)*100 As Growth2017,
    (((SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 1 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 1 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END)) +
    ((SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END)) +

```

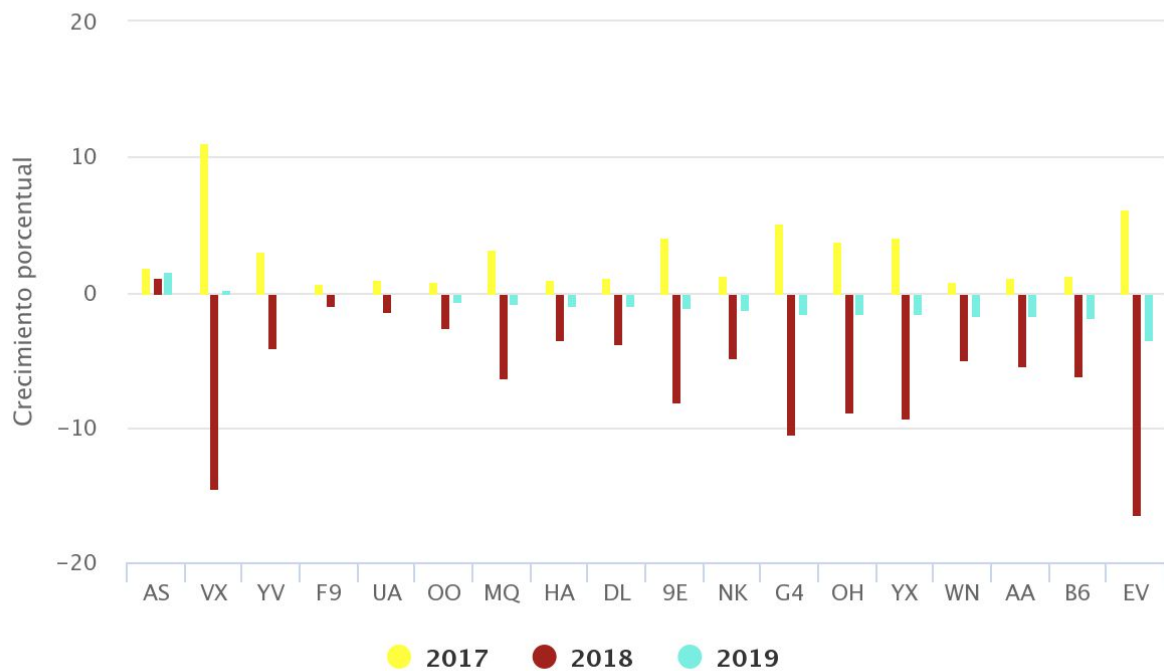
```

((SUM(CASE WHEN Quarter = 4 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))/3)*100 As Growth2018,
((((SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 1 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 1 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 2 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 3 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 1 and "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 4 and "Year" = '2017' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 1 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 1 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 2 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END)) +
((SUM(CASE WHEN Quarter = 4 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0 END) -
SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END))*1.0/SUM(CASE WHEN Quarter = 3 AND "Year" = '2018' NOT NULL THEN 1 ELSE 0
END)))/7)*100 as CrecimientoPromedio
FROM ontime ontime_externo
GROUP BY Reporting_Airline
ORDER BY CrecimientoPromedio DESC

```

Reporting_Airline	Growth2017	Growth2018	CrecimientoPromedio
AS	1.855016208	1.117236375	1.538824851
VX	11.10625853	-14.44372001	0.156267727
YV	3.06310772	-4.111738576	-0.011826407
F9	0.65739326	-0.934171843	-0.02470607
UA	0.927987689	-1.385551106	-0.063528937
OO	0.840972863	-2.678687533	-0.667453021
MQ	3.188523943	-6.343825914	-0.896768853
HA	0.939780885	-3.547460015	-0.983322358
DL	1.123864105	-3.858671298	-1.01150821
9E	4.079972628	-8.109481792	-1.144079267
NK	1.252037193	-4.781143558	-1.3336117
G4	5.061545486	-10.53044318	-1.620735371
OH	3.743420275	-8.785742404	-1.626220873
YX	4.074878765	-9.241544295	-1.63215969
WN	0.887988113	-5.035555693	-1.650673518
AA	1.069188225	-5.452445718	-1.725797751
B6	1.20714574	-6.100411804	-1.924664636
EV	6.198797537	-16.41387158	-3.49234637

## Crecimiento a lo largo del año



### 8. Octavo Query

- Para obtener cuales aerolíneas tienen la mayor cantidad de vuelos en los estados más poblados (es decir, los estados con la mayor cantidad de aeropuertos)

```
SELECT Reporting_Airline, COUNT(DISTINCT origen) as estados
```

```

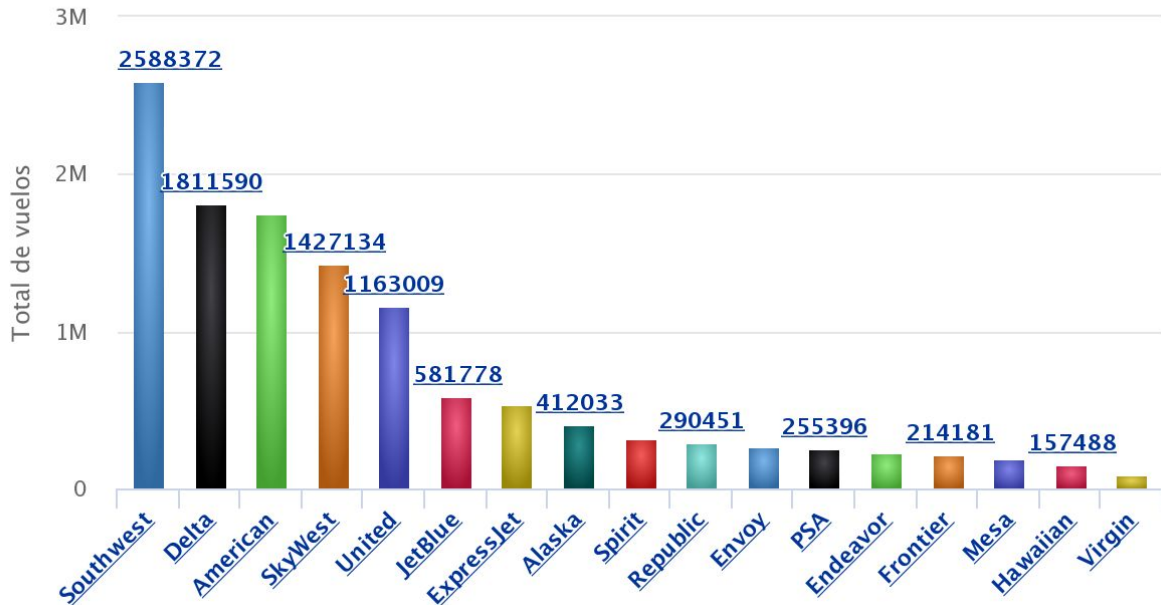
FROM ontime o
JOIN (
    SELECT COUNT(DISTINCT DestAirportID) as CantAeropuertos, OriginStateNm as origen
    FROM ontime
    GROUP BY OriginStateNm
    ORDER BY CantAeropuertos DESC
    LIMIT 10
) tabla ON o.DestStateNm = tabla.origen
GROUP BY Reporting_Airline
ORDER BY estados DESC

```

Reporting_Airline	Estados
AA	10
AS	10
B6	10
DL	10
F9	10
G4	10
NK	10
OO	10
UA	10
WN	10
YV	10
EV	9
9E	8
MQ	8
YX	8
OH	7
VX	6
HA	2

## Cantidad de vuelos a los 10 estados más poblados. 2017–2018

Click a las columnas para ver la cantidad de vuelos por estado más poblado.



- b. Para obtener el desglose de cuántos vuelos en cada estado (de los más poblados) tiene cada aerolínea

```
SELECT unique_carriers.Description, SUM(CASE WHEN DestStateAbr = "TX" THEN 1 ELSE 0
END) as texas,
      SUM(CASE WHEN DestStateAbr = "CA" THEN 1 ELSE 0 END) as california,
      SUM(CASE WHEN DestStateAbr = "FL" THEN 1 ELSE 0 END) as Florida,
      SUM(CASE WHEN DestStateAbr = "AK" THEN 1 ELSE 0 END) as Alaska,
      SUM(CASE WHEN DestStateAbr = "MI" THEN 1 ELSE 0 END) as Michigan,
      SUM(CASE WHEN DestStateAbr = "NY" THEN 1 ELSE 0 END) as NY,
      SUM(CASE WHEN DestStateAbr = "CO" THEN 1 ELSE 0 END) as Colorado,
      SUM(CASE WHEN DestStateAbr = "IL" THEN 1 ELSE 0 END) as Illinois,
      SUM(CASE WHEN DestStateAbr = "NC" THEN 1 ELSE 0 END) as North_Carolina,
      SUM(CASE WHEN DestStateAbr = "VA" THEN 1 ELSE 0 END) as Virginia
FROM ontime
LEFT JOIN unique_carriers ON ontime.Reporting_Airline = unique_carriers.Code
GROUP BY Reporting_Airline
```



Description	Texas	California	Florida	Alaska	Michigan	NY	Colorado	Illinois	North_Carolina	Virginia
Endeavor Air Inc.	4842	0	5331	0	20133	58208	28	2827	12686	8996
American Airlines Inc.	323815	158603	197394	536	11861	81659	26073	123171	207530	61256
Alaska Airlines Inc.	8062	98076	4626	64221	1242	4069	3426	5020	942	4802
JetBlue Airways	7160	61354	129253	424	2722	119592	2220	4808	8980	29663
Delta Air Lines Inc.	47988	127628	158706	5122	117602	116833	22323	21020	33997	36818
ExpressJet Airlines Inc.	127747	0	10710	0	13641	29996	954	36732	13900	21615
Frontier Airlines Inc.	11380	12573	35100	0	2556	5285	45532	5716	5514	4348
Allegiant Air	1730	4084	25295	0	1509	2223	343	2980	2832	774
Hawaiian Airlines Inc.	0	9913	0	0	0	693	0	0	0	0
Envoy Air	62247	0	15385	0	6203	23477	1000	59634	9946	7213
Spirit Air Lines	29087	26629	70483	0	19088	9294	7121	18742	198	522
PSA Airlines Inc.	234	0	9403	0	2307	7454	0	2738	104619	32926
SkyWest Airlines Inc.	44173	254080	2148	374	109150	23445	120623	164143	14172	14169
United Air Lines Inc.	150601	216921	71881	1987	4755	21426	122110	145456	10353	58729
Virgin America	5684	49695	2540	0	0	6107	1215	1713	164	3497
Southwest Airlines Co.	331169	464969	216914	0	19342	54337	132526	159992	25762	41142
Mesa Airlines Inc.	80297	9045	3170	0	2276	2830	2384	1803	2942	23448
Republic Airline	20332	0	18657	0	8878	40125	4550	16232	15939	32743

## 9. Noveno Query

```

SELECT Reporting_Airline, RetrasosAerolineas, SUM(Distance)/(SUM(AirTime)/60) as
Velocidad
FROM ontime
JOIN (
    SELECT SUM(arrDelay) as RetrasosAerolineas, Reporting_Airline as aero
    FROM ontime
    WHERE arrDelay>0
    GROUP BY Reporting_Airline
    ORDER BY RetrasosAerolineas ASC
) Delays ON ontime.Reporting_Airline = Delays.aero
GROUP BY Reporting_Airline
ORDER BY RetrasosAerolineas ASC

```

Reporting_Airline	RetrasosAerolíneas	Velocidad
HA	934952	448.9836882
VX	1490442	474.8476392
G4	1560343	452.8840935
YV	3010744	413.0648549
9E	3319270	396.6535583
MQ	3414205	390.4106699
YX	3642530	408.2470841
AS	3722874	461.9403788
OH	3766518	374.0287435
F9	4039137	459.7503417
NK	4664065	459.4964097
EV	9170694	393.2639742
B6	11365918	453.7089659
UA	15558135	465.2724202
DL	18638450	445.8065021
OO	21762259	392.2127397
AA	22526555	451.9282106
WN	28443506	439.398118

## 10. Décimo Query

```

SELECT Reporting_Airline as Aerolineas,
SUM(CASE WHEN DistanceGroup = 1 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion1,
SUM(CASE WHEN DistanceGroup = 2 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion2,
SUM(CASE WHEN DistanceGroup = 3 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion3,
SUM(CASE WHEN DistanceGroup = 4 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion4,
SUM(CASE WHEN DistanceGroup = 5 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion5,
SUM(CASE WHEN DistanceGroup = 6 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion6,
SUM(CASE WHEN DistanceGroup = 7 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion7,
SUM(CASE WHEN DistanceGroup = 8 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion8,
SUM(CASE WHEN DistanceGroup = 9 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion9,
SUM(CASE WHEN DistanceGroup = 10 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion10,

```

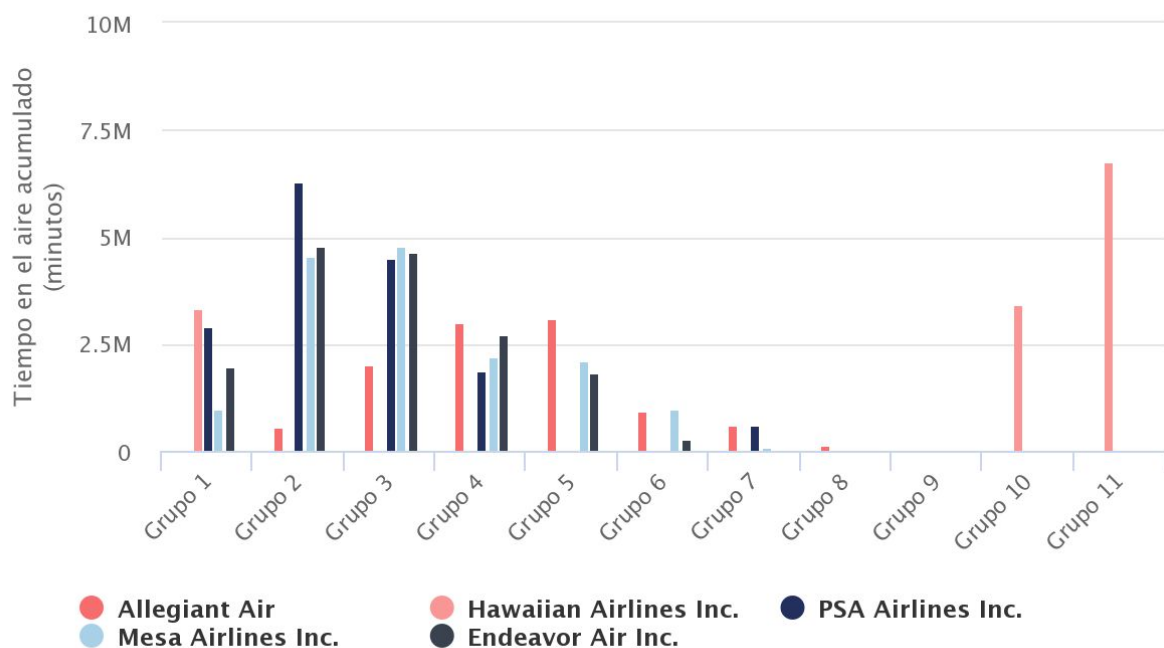
```

SUM(CASE WHEN DistanceGroup = 11 THEN AirTime ELSE 0 END) as
TiempoEnAireClasificacion11
FROM ontime
GROUP BY Reporting_Airline
ORDER BY (TiempoEnAireClasificacion1 + TiempoEnAireClasificacion2 +
TiempoEnAireClasificacion3 +
TiempoEnAireClasificacion4 + TiempoEnAireClasificacion5 +
TiempoEnAireClasificacion6 +
TiempoEnAireClasificacion7 + TiempoEnAireClasificacion8 +
TiempoEnAireClasificacion9 +
TiempoEnAireClasificacion10 + TiempoEnAireClasificacion11) ASC

```

Aerolíneas	TiempoEnAireClasificacion1	TiempoEnAireClasificacion2	TiempoEnAireClasificacion3	TiempoEnAireClasificacion4	TiempoEnAireClasificacion5	TiempoEnAireClasificacion6	TiempoEnAireClasificacion7	TiempoEnAireClasificacion8	TiempoEnAireClasificacion9	TiempoEnAireClasificacion10	TiempoEnAireClasificacion11
G4	105	585471	2000750	3001390	3085820	940982	609886	121235	0	0	0
HA	3331268	62080	0	0	0	0	0	0	0	3423027	6740616
OH	2895698	6301555	4514690	1894394	66073	16010	0	0	0	0	0
YV	972885	4565296	4789459	2199722	2122559	984255	74222	0	0	0	0
9E	1951695	4805198	4652937	2703585	1843842	259586	0	0	0	0	0
VX	159349	1347185	538790	638676	889991	719215	777581	695367	505506	5965932	4490336
MQ	2291606	6300937	4710183	3725112	805700	170410	0	0	0	0	0
YX	1935157	5593907	6071947	5061868	4388133	1927236	392137	0	0	0	0
F9	302	1054636	2853378	7405083	6110223	3284985	4806881	1584823	2250867	277476	0
EV	3756796	12287878	12712179	7436186	1672422	178	0	0	0	0	0
NK	313184	1916305	4434689	10490420	10552847	3301698	4514995	2212088	2347262	2308462	89809
AS	955211	1497224	6537352	10553774	4402636	5279107	5605352	3203467	1634584	15019104	13684600
B6	1872557	6733801	4190359	11570283	23306378	770228	5892567	1557627	3334698	11595220	10242042
OO	13646281	28729456	26881864	15091943	8084995	7319361	3502888	80993	0	0	0
UA	2488227	8729258	15417023	26209918	17430939	16287976	24048868	10826525	7705144	28099901	20506917
DL	5931808	22702091	41534583	29007196	23049872	12865391	23161873	18261609	15741731	10017000	12075061
AA	5915986	13278483	31096731	33515835	41591695	23111888	20328392	14267029	14938066	21890538	17107835
WN	7474725	48854894	41017743	55668770	38675487	20250947	24376926	16696183	6179544	4798617	183230

## Tiempo en el aire de cada aerolínea respecto a la distancia recorrida total



## 11. Puntuación o peso de cada query para cada aerolínea

Se decidió que cada pregunta tendría un peso o punteo para así poder decidir, de acuerdo a estos punteos, cuál aerolínea se consideraba tuvo el mejor desempeño en los años 2017 y 2018. El peso de cada pregunta quedó de la siguiente manera:

- Pregunta 1: 2pts
- Pregunta 2: 1pts
- Pregunta 3: 3pts
- Pregunta 4: 4pts
- Pregunta 5: 2pts
- Pregunta 6: 3pts
- Pregunta 7: 4pts
- Pregunta 8: 2pts
- Pregunta 9: 1pt
- Pregunta 10: 2pts

De acuerdo con esto, los resultados globales quedaron de la siguiente manera:

Nombre	Codigo	Pregunta 1	Pregunta 2	Pregunta 3	Pregunta 4	Pregunta 5	Pregunta 6	Pregunta 7	Pregunta 8	Pregunta 9	Pregunta 10	Resultados
Alaska Airlines Inc.	A5							4				4
Allegiant Air	G4	2	1		4	2				1	2	12
American Airlines Inc.	AA		1	3					2			6
Delta Air Lines Inc.	DL		1	3					2			6
Endeavor Air Inc.	9E	2				2	3			1	2	10
Envoy Air	MQ				4		3					7
ExpressJet Airlines Inc.	EV						3					3
Frontier Airlines Inc.	F9							4				4
Hawaiian Airlines Inc.	HA	2			4	2				1	2	11
JetBlue Airways	B6											0
Mesa Airlines Inc.	YV							4		1	2	7
PSA Airlines Inc.	OH						3				2	5
Republic Airline	YX				4							4
SkyWest Airlines Inc.	OO	2	1	3		2	3		2			13
Southwest Airlines Co.	WN			3					2			5
Spirit Air Lines	NK											0
United Air Lines Inc.	UA		1	3				4	2			10
Virgin America	VX	2			4	2		4		1		13

## C. Discusión de Resultados

Para realizar el análisis de cada pregunta, se le asignó un peso (puntuación) a cada pregunta en base a la importancia que tenían los resultados de estas para evaluar cuál era la mejor aerolínea. En cuanto a esto, se evaluaron unas características esenciales para medir cual era la mejor aerolínea como por ejemplo: vuelos cancelados, cantidad de tiempo de atrasos, proyección a futuro, eficiencia, seguridad y experiencia. Para todas las preguntas se hizo una

clasificación de las cinco con mejor resultado y únicamente a esas aerolíneas se les asignó el peso de la pregunta. Al terminar de asignar el peso, se realizó la sumatoria de cuántos puntos había obtenido cada aerolínea.

En cuanto a la primer pregunta, se comenzó con un atributo básico pues es importante saber cuántos vuelos ha cancelado esta aerolínea, en este caso solo se tomaron en cuenta los vuelos que fueron cancelados por motivos de la misma aerolínea. Dado a que no es por motivos meteorológicos o del aeropuerto, el vuelo pudo haber sido cancelado por un desperfecto mecánico en el avión que iba a realizar este vuelo. Si bien el vuelo se cancela por seguridad del pasajero, se puede tener un indicio en cuanto a la organización y chequeo mecánico continuo de sus aviones.

A continuación analizaremos la pregunta 2 y la pregunta 8 ya que ambas se basan en los estados hacia los que vuelan las aerolíneas. En la pregunta 2 se obtiene únicamente el resultado de a cuantos estados dentro de Estados Unidos viajan las aerolíneas, en esta pregunta nos interesaron las aerolíneas que viajan a todos los estados o en su defecto a la mayoría de los 51 (según la base de datos). Y en cuanto a la pregunta 8 nos interesó la cantidad de vuelos por cada estado más poblado. La manera en que se consideró cuales eran los estados más poblados es en base a la cantidad de sus aeropuertos, esto pues entre más aeropuertos tenga el estado mayor será su concurrencia de vuelos tanto nacionales como internacionales, de igual forma da indicio si el estado es una de los más grandes o no de Estados Unidos. En cuanto a estos resultados se pudo observar que habían aerolíneas que viajaban a su mayoría de estados poblados pero la cantidad de sus vuelos era mayor en todos aquellos estados que eran más turísticos, entiéndase California o Florida a comparación de Texas. En cuanto a esta información las aerolíneas que viajan a distintos destinos puede reflejar un servicio de mayor calidad y mayor experiencia. Dado a que al viajar a más lugares tienen que ser más cuidadosos en cuanto a la seguridad del equipaje (que no se pierda ninguna maleta), de igual forma se puede inferir que sus aviones permiten una mayor cantidad de pasajeros.

Para las preguntas 3 y 9, se decidió analizar los tiempos de retraso. En el caso de la pregunta 3, se tomó el tiempo de *no* retraso, es decir, analizar cuánto tiempo logró ahorrar cada

aerolínea para llegar a su destino, cuando habían salido a tiempo de su origen; esta variable fue tomada en consideración ya que es beneficioso que un vuelo llegue antes de tiempo y así poder utilizar este tiempo ahorrado para otros asuntos, como el mantenimiento y limpieza del avión, además que deja una impresión positiva en los pasajeros y, por tanto, en futuros clientes. Para los resultados de la pregunta 4 se analizó las aerolíneas que habiendo salido a tiempo llegaron antes a su destino, esto nos da un indicio de la experiencia del piloto y su eficiencia en el aire, las aerolíneas que más vuelos tengan con estas características es importante tanto como para el usuario como para el piloto y demás trabajadores de la aerolínea. Ya que este delta de tiempo entre la hora de llegada antes y la hora de llegada programada, permite que los pasajeros del vuelo tengan más tiempo para acomodarse, salir del aeropuerto. En cuanto a los trabajadores, esto implica un leve descanso entre sus labores, lo cual es importante ya que al estar descansados pueden desempeñarse mejor en su trabajo, brindando un servicio de calidad.

De la misma manera, en la pregunta 9, se clasificó cuáles aerolíneas tienen el menor tiempo de retraso - porque cualquier retraso puede significar disgustos en los pasajeros o hasta problemas dentro de las mismas aerolíneas, ya que en el caso que ese vuelo en específico sea una escala o conexión, atrasaría al siguiente vuelo y le daría menos tiempo a los pasajeros de movilizarse entre puerta y puerta del aeropuerto.

Seguidamente se analizó la pregunta 5, en la cual se encontró la frecuencia de vuelos para las aerolíneas que tienen menor cantidad de vuelos. Esto da indicio de que tan seguido una aerolínea realiza un vuelo, para el usuario final es importante pues si los servicios son buenos y son de las pocas aerolíneas que vuelan a un cierto destino, el realizar vuelos consecutivamente le da prestigio a esta aerolínea.

Para la pregunta 6 y 10, se analizó la distancia recorrida acumulada de todos los vuelos de una aerolínea, y, en el caso de la 6, respecto a la cantidad de vuelos que ha tenido. Para la primera, se hizo de esta manera debido a que se quería analizar las millas totales que cada aerolínea tenía acumuladas de sus vuelos, pero no era lo mismo que una aerolínea recorriera grandes distancias, pero tuviera pocos vuelos, a una aerolínea que recorriera grandes distancias y lo hiciera con múltiples vuelos. Por tanto, se decidió hacer una razón entre las

dos variables, y ordenar las aerolíneas de esta manera. Para la pregunta 10, se decidió analizar, además de la distancia, el tiempo que un avión se mantenía en el aire, debido a que, podía ser que múltiples vuelos recorrieran las mismas distancias, pero algunos se tardan menos, lo que lleva a un adelanto o atraso de llegada, que es importante por las mismas razones planteadas para las preguntas 3 y 9.

En cuanto a la pregunta 7, esta se consideró de las preguntas más relevantes y de las más influyentes, ya que no solo se analizan los datos que se tienen en el presente, sino que se intenta hacer una proyección a futuro sobre las operaciones de cada aerolínea. De esta manera, podríamos saber si las operaciones de una aerolínea crecerán o decrecerán en un futuro y cuál será la que tendrá el mayor crecimiento. Esta pregunta se planeó para obtener una visión desde una persona que decida invertir en una aerolínea, claramente la aerolínea que tenga un mayor crecimiento a futuro será la mejor para invertir a comparación a la que ha decreciendo en cuanto a sus operaciones. La manera en la que se obtuvo dicha proyección fue: se tomaron las operaciones - cantidad de vuelos - por aerolínea por cuartos de un año - tres meses - y se obtuvo la tasa de crecimiento de un cuarto, respecto al siguiente cuarto, de modo que:  $\frac{\text{Segundo Cuarto} - \text{Primer Cuarto}}{\text{Primer Cuarto}}$ , y de esta manera se realizaron todos los cuartos del año, y se hizo un promedio entre ellos, para cada año.

#### D. Conclusión

De acuerdo a los resultados obtenidos en la tabla II.B.11, las aerolíneas con los mejores puntajes fueron SkyWest Airlines Inc. y Virgin America, ambas con 13 puntos. Para decidir cuál de las dos sería la aerolínea con mayor desempeño, se tomó en consideración la cantidad de vuelos por años y su crecimiento en operaciones (pregunta 7). SkyWest tuvo, en 2017, 717,674 vuelos en total y para 2018, 709,460; sin embargo, para el 2019, tiene predicho un decrecimiento en sus operaciones de 0.67%. Por otro lado, Virgin America tuvo un total de 71,658 vuelos, pero para el 2018 decayó enormemente, con un total de 17,670 vuelos; a pesar de esto, esta aerolínea tiene predicha un crecimiento de 0.15% en sus operaciones. Por tanto, debido a que Virgin America tuvo una variación considerablemente grande entre sus

operaciones del 2017 y 2018, se decidió que dichas operaciones serían muy impredecibles o variantes para el futuro y probablemente su predicción no era tanto de fiar. Por otro lado, SkyWest se mantuvo relativamente constante en sus operaciones y a pesar que su predicción fuera negativa, se consideró que no es lo suficiente grande como para ser una gran desventaja, por tanto, se concluyó que la aerolínea con mejor desempeño en los años 2017 y 2018 fue Skywest Airlines Inc.