**Text

Description automatically generated with medium confidenceText

Description automatically generated with medium confidence**

**Report title: Apache Pig Weekly Project - Airlines Delay**

**Group name: Ninjaz**

**Dataset description:**

|  |  |  |
| --- | --- | --- |
| **no.** | **feature** | **discription** |
| 1 | Year | 2008 |
| 2 | Month | 1-Dec |
| 3 | DayofMonth | Jan-31 |
| 4 | DayOfWeek | 1 (Monday) - 7 (Sunday) |
| 5 | DepTime | actual departure time (local, hhmm) |
| 6 | CRSDepTime | scheduled departure time (local, hhmm) |
| 7 | ArrTime | actual arrival time (local, hhmm) |
| 8 | CRSArrTime | scheduled arrival time (local, hhmm) |
| 9 | UniqueCarrier | 1. unique carrier code |
| 10 | FlightNum | flight number |
| 11 | TailNum | plane tail number: aircraft registration, unique aircraft identifier |
| 12 | ActualElapsedTime | in minutes |
| 13 | CRSElapsedTime | in minutes |
| 14 | AirTime | in minutes |
| 15 | ArrDelay | arrival delay, in minutes: A flight is counted as "on time" if it operated less than 15 minutes later the scheduled time shown in the carriers' Computerized Reservations Systems (CRS). |
| 16 | DepDelay | departure delay, in minutes |
| 17 | Origin | origin IATA airport code |
| 18 | Dest | destination IATA airport code |
| 19 | Distance | in miles |
| 20 | TaxiIn | taxi in time, in minutes |
| 21 | TaxiOut | taxi out time in minutes |
| 22 | Cancelled | \*was the flight cancelled |
| 23 | CancellationCode | reason for cancellation (A = carrier, B = weather, C = NAS, D = security) |
| 24 | Diverted | 1 = yes, 0 = no |
| 25 | CarrierDelay | in minutes: Carrier delay is within the control of the air carrier. Examples of occurrences that may determine carrier delay are: aircraft cleaning, aircraft damage, awaiting the arrival of connecting passengers or crew, baggage, bird strike, cargo loading, catering, computer, outage-carrier equipment, crew legality (pilot or attendant rest), damage by hazardous goods, engineering inspection, fueling, handling disabled passengers, late crew, lavatory servicing, maintenance, oversales, potable water servicing, removal of unruly passenger, slow boarding or seating, stowing carry-on baggage, weight and balance delays. |
| 26 | WeatherDelay | in minutes: Weather delay is caused by extreme or hazardous weather conditions that are forecasted or manifest themselves on point of departure, enroute, or on point of arrival. |
| 27 | NASDelay | in minutes: Delay that is within the control of the National Airspace System (NAS) may include: non-extreme weather conditions, airport operations, heavy traffic volume, air traffic control, etc. |
| 28 | SecurityDelay | in minutes: Security delay is caused by evacuation of a terminal or concourse, re-boarding of aircraft because of security breach, inoperative screening equipment and/or long lines in excess of 29 minutes at screening areas. |
| 29 | LateAircraftDelay | in minutes: Arrival delay at an airport due to the late arrival of the same aircraft at a previous airport. The ripple effect of an earlier delay at downstream airports is referred to as delay propagation. |

Graphical user interface, text, application

Description automatically generated**Q1:** How many minutes the arrival delay for each flight with distance greater than 500 miles?

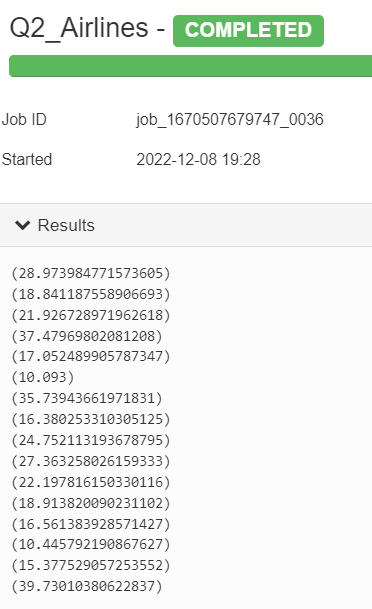
Figure / Question one (script result)

**Output:**

Illustrated in figure 1.

**Observation:**

we noticed that the minimum arrival delay starts with 15 minutes.



**Q2:** What is the average minute delay for each carrier?

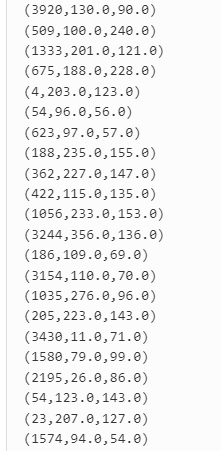
**Output:**

Illustrated in figure 2.

**Observation:**

The average delay ranges from 10 to 40 minutes.

Figure / Question two (script result)

**Q3:** What is the difference between the arrival and departure time of the flights and the actual elapsed time for the flight?

**Output:**

Illustrated in figure 3.

**Observation:** If we compare the Elapsed time of the flight and the actual time of the flight we can see that a lot of the flights had an extreme amount of delay.

Figure / Question three (script result)

**Q4:** How many flights are departing from Indiana airport to Baltimore airport?

**Output:**

Illustrated in figure 4.

**Observation:**

We have about 9 flights departing from Indiana airport to Baltimore airport.

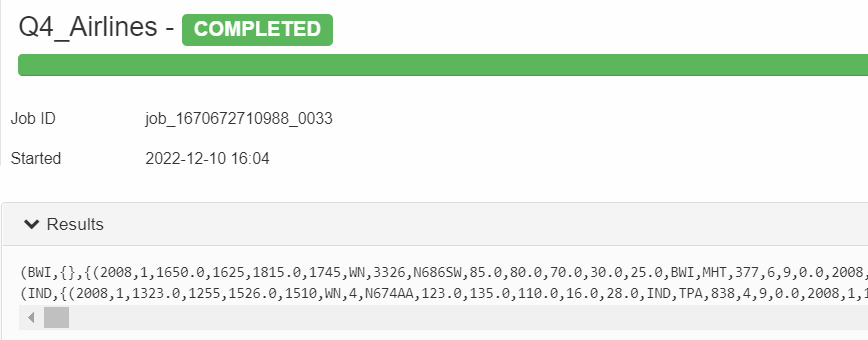


Figure / Question four (script result)

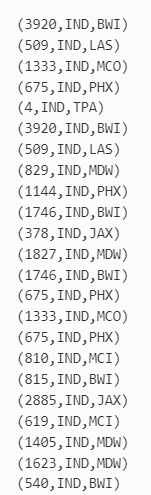
**Q5:** What are the flights coming from Indiana Airport?

Figure / Question five (script result)

**Output:**

Illustrated in figure 5.

**Observation:**

We have 23 flights coming from Indiana airport only.

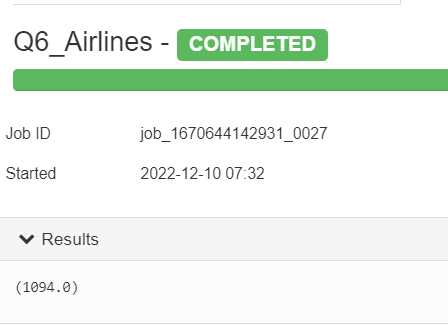
**Q6:** What is the maximum delay?

Figure / Question Six (script result)

**Output:**

Illustrated in figure 6.

**Observation:**

The maximum delay of a flight was 1094 min.

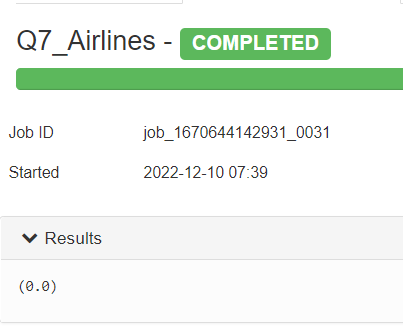
**Q7:** What is the minimum delay?

Figure / Question seven (script result)

**Output:**

Illustrated in figure 7.

**Observation:**

The minimum record for delay is 0.0

Graphical user interface, application

Description automatically generated

**Q8:** Count the number flights of each unique carriers.

**Output:**

Illustrated in figure 8.

**Observation:**

We have number of flights for each carrier more than 4000 flights , with the southwest airlines being the highest and Hawaii airlines as the lowest.

Figure / Question eight (script result)

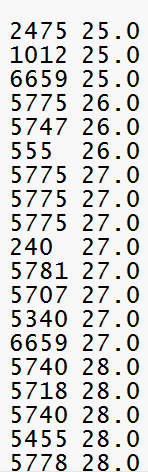
**Q9:** What is the Actual elapsed time for flights in ascending order?

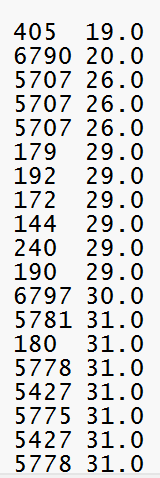
Figure /Question nine (script result)

**Output:**

Illustrated in figure 9.

**Observation:**

We can see that the actual elapsed time ranges from 25 to 28 minutes.

****

**Q10:** What is the CRSelapsed time for flights in ascending order?

Figure / Question ten (script result)

**Output:**

Illustrated in figure 10.

**Observation:**

We can see that the CRS elapsed time ranges from 20to 30 minutes.