

Impact of weather on hotel booking cancellation

Presented by Arlan Amangeldin

Outline

Business use case



Dataset overview



Tool stack



Price estimation





Business use case



Hotel chains lose significant revenue from cancellations.

Looking on how weather impacts hotel bookings helps to:

- Predict cancellations based on weather forecasts
- Optimize pricing strategies during poor weather periods
- Adjust workforce according to expected arrivals
- Improve marketing campaigns targeting unpleasant weather condition periods to bring more customers

Stakeholders: Revenue manager, Operations director, Hotel Owners/Investors

Dataset: Hotel Booking Demand

Kaggle dataset contains information on hotels bookings for two types of hotels in two different cities of Portugal.

Resort hotel is located in Algarve, whereas city hotel is located in Lisbon.

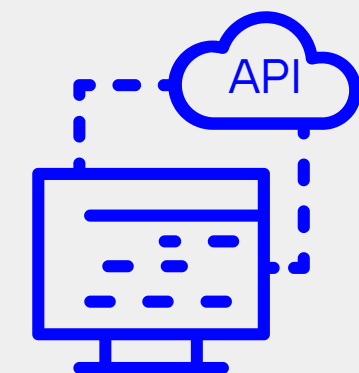
Key columns include measurements like cancelled flag and average day revenue. This will be used for quantifying of success.

API: Open-Meteo Weather

Weather metrics for temperature, precipitation, rain, snow, wind and sun are derived through API call using Python script.

Open-Meteo API is a best choice as it does not require any method of authorization and has a high requests limit per day.

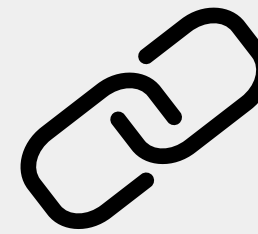
Data pipeline



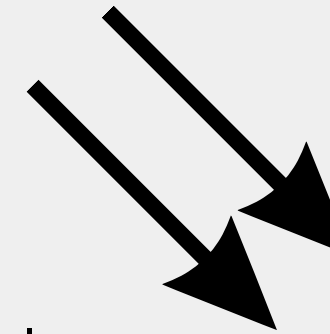
**Open-Meteo
API**



AWS Lambda -
To fetch historical
temperature values
and other relevant
measurments



AWS Eventbridge -
To update weather
information on a
weekly basis

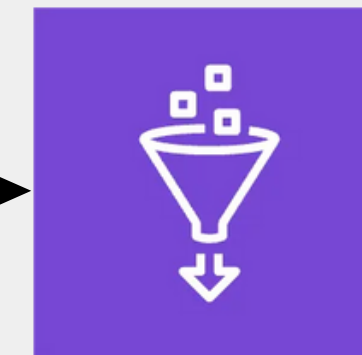
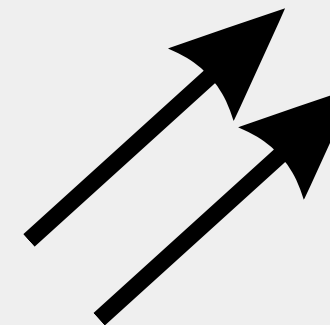


S3

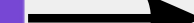
**S3 bucket
(data
storage)**
To store data
from Lambda
and
Cloudshell
services
(dataset data
and data from
API)



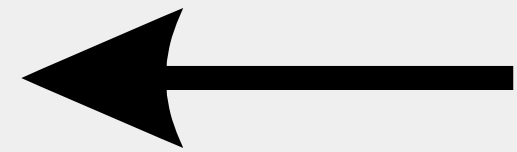
Load hotel booking
dataset from Kaggle
into S3 bucket folder



AWS Glue -
To combine
data from both
CSV and API



AWS Athena
- Analyze how
temperature
and other
weather
measurement
s affect hotels
revenue and
booking
cancelations



This setup lets stakeholders run queries on booking and weather data together, so they can spot patterns and make decisions on pricing, staffing, and marketing based on weather forecasts.

S3 bucket

S3 bucket is created with following folder structure:

General purpose buckets

All AWS Regions

Directory buckets

General purpose buckets (1) [Info](#)

Copy ARN

Empty

Delete

Create bucket

Find buckets by name

< 1 >

Name

▲

AWS Region

▼

Creation date

▼

[hotel-weather-analysis-sgmxrq](#)

US East (N. Virginia) us-east-1

January 9, 2026, 22:34:26 (UTC+01:00)

hotel-weather-analysis-sgmxrq [Info](#)

Objects

Metadata

Properties

Permissions

Metrics

Management

Access Points

Objects (3)

Copy S3 URI

Copy URL

Download

Open [↗](#)

Delete

Actions ▼

Create folder

Upload

Find objects by prefix

< 1 >

Name

▲

Type

▼

Last modified

▼

Size

▼

Storage class

▼

[athena-results/](#)

Folder

-

-

-

[processed-data/](#)

Folder

-

-

-

[raw-data/](#)

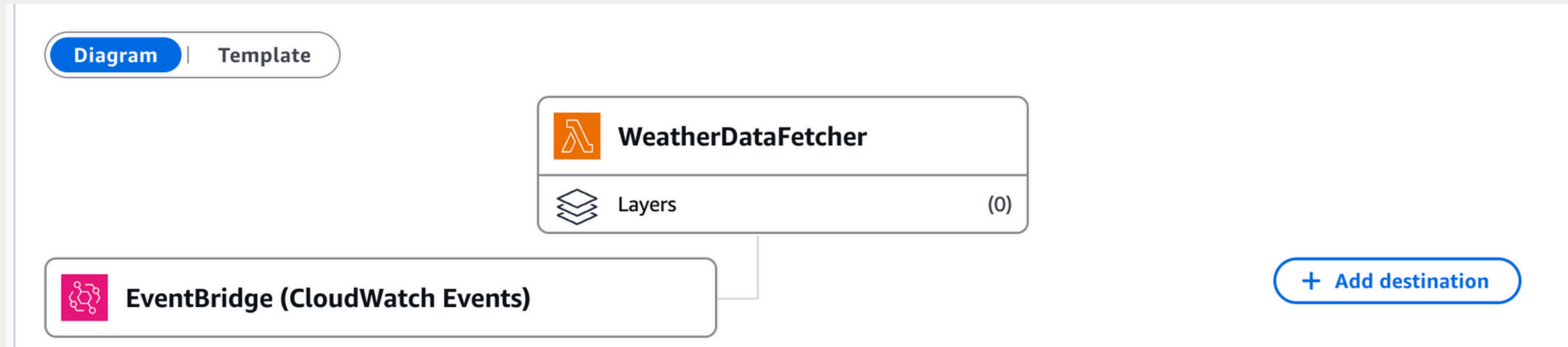
Folder

-

-

-

Lambda and EventBridge



Lambda function to extract weather data was created and information is updated each 7 days as triggered by EventBridge rule

AWS Glue and Crawler

1.

Databases (2) Last updated (UTC) January 12, 2026 at 00:44:41 Edit Delete Add database

A database is a set of associated table definitions, organized into a logical group.

| <input type="checkbox"/> | Name | Description | Location URI | Source catalog | Created on (UTC) |
|--------------------------|----------------------------------|-----------------------------------------|---------------------------|----------------|------------------------------|
| <input type="checkbox"/> | default | default database | file:/tmp/spark-warehouse | 807214179386 | January 10, 2026 at 23:48:24 |
| <input type="checkbox"/> | hotel_weather_db | Database for hotel bookings and weather | - | 807214179386 | January 9, 2026 at 23:28:05 |

Crawlers (3) Info Last updated (UTC) January 12, 2026 at 00:44:16 Action Run Create crawler

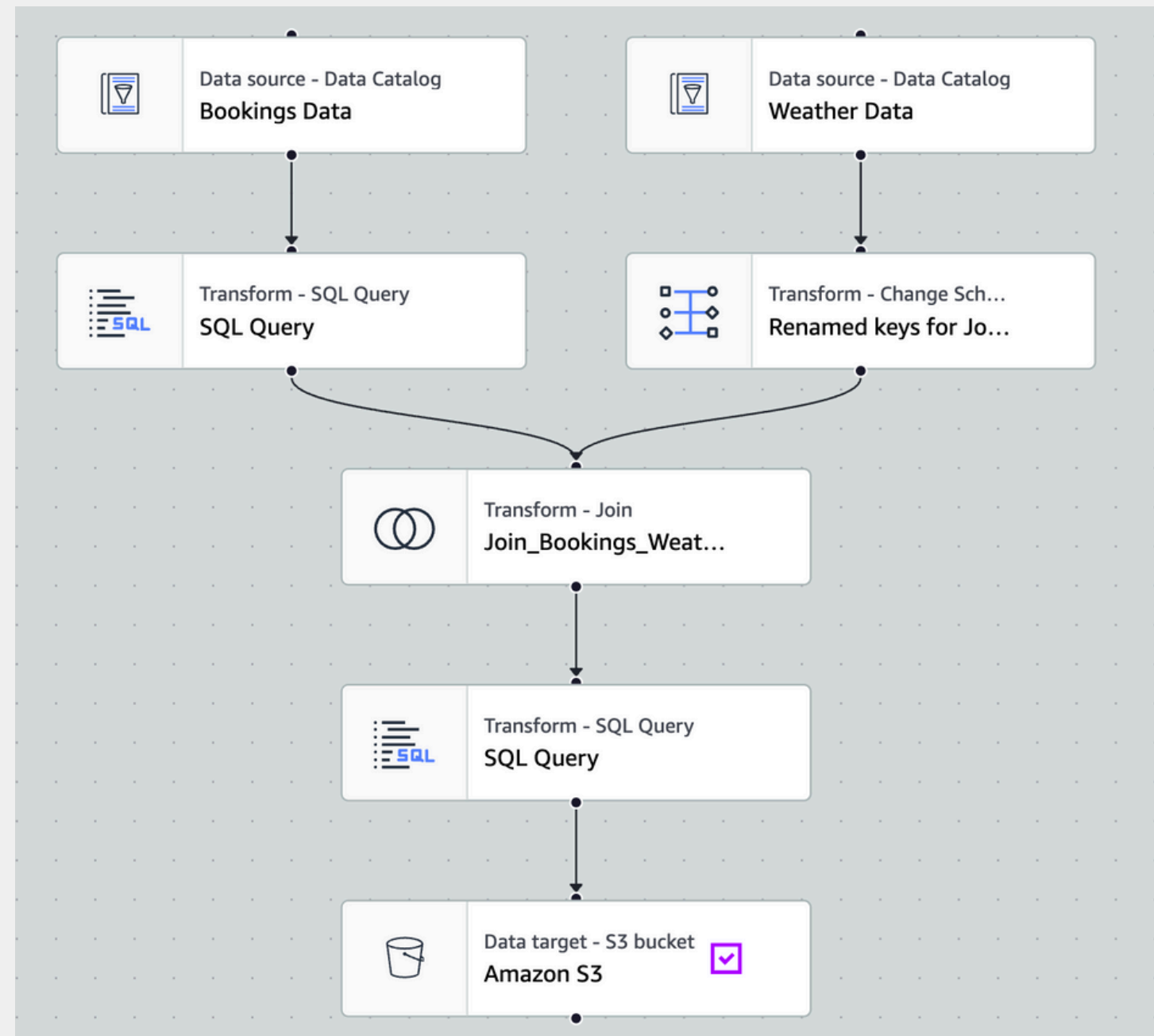
View and manage all available crawlers.

| <input type="checkbox"/> | Name | State | Schedule | Last run | Last run timestamp | Log | Table changes from ... |
|--------------------------|----------------------------------------|-------|----------|-----------|--------------------------|--------------------------|------------------------|
| <input type="checkbox"/> | combined-data-crawler | Ready | | Succeeded | January 11, 2026 at 2... | View log | 1 created |
| <input type="checkbox"/> | hotel-bookings-crawler | Ready | | Succeeded | January 10, 2026 at 2... | View log | 1 created |
| <input type="checkbox"/> | weather-data-crawler | Ready | | Succeeded | January 10, 2026 at 2... | View log | 1 created |

2.

Glue database was created and crawlers for booking, weather and combined data were configured

ETL job with Glue



Visual ETL was created to combine data from Bookings and Weather sources. Certain new fields were created and final combined table is flattened out (to remove struct type objects)

KPI

Weather impact on cancellation rate and revenue

Lead time against cancellation risk

Guest composition and weather preference

Weather impact on cancellation rate

| hotel | temp_category | total_bookings | cancellations | cancellation_rate | hotel | wind_level | bookings | avg_wind_speed | cancel_rate | avg_rate |
|--------------|--------------------|----------------|---------------|-------------------|--------------|-----------------------|----------|----------------|-------------|----------|
| City Hotel | Moderate (15-25°C) | 51780 | 22241 | 42.95 | City Hotel | Strong (>40 km/h) | 955 | 43.6 | 43.14 | 81.03 |
| City Hotel | Cold (<15°C) | 24572 | 9781 | 39.81 | City Hotel | Calm (<20 km/h) | 30270 | 16,1 | 42.41 | 106.82 |
| Resort Hotel | Moderate (15-25°C) | 25547 | 7283 | 28.51 | City Hotel | Moderate (20-40 km/h) | 46897 | 26,2 | 42.06 | 107.54 |
| Resort Hotel | Warm (>25°C) | 5711 | 1970 | 34.49 | Resort Hotel | Moderate (20-40 km/h) | 27408 | 27,5 | 28.22 | 100.28 |
| Resort Hotel | Cold (<15°C) | 8802 | 1869 | 21.23 | Resort Hotel | Calm (<20 km/h) | 10473 | 15,9 | 27.81 | 92.05 |
| City Hotel | Warm (>25°C) | 2978 | 1080 | 36.27 | Resort Hotel | Strong (>40 km/h) | 1427 | 44.7 | 27,4 | 63.93 |

City Hotel cancellation rates stay pretty much the same regardless of wind or temperature. Resort Hotel is more affected — prices drop hard when it's windy (\$64 vs \$100). Cold weather at the resort actually has lowest cancellations (21%) which is weird for a beach place

| hotel | rain_category | bookings | avg_precipitation | cancel_rate | avg_rate | revenue |
|--------------|------------------------|----------|-------------------|-------------|----------|------------|
| City Hotel | Dry (0mm) | 54097 | 0.0 | 42.11 | 109.62 | 3474726.30 |
| City Hotel | Light Rain (0-5mm) | 16185 | 1,2 | 42.37 | 102.07 | 965201.04 |
| City Hotel | Moderate Rain (5-15mm) | 6173 | 8,4 | 44.45 | 95.47 | 331140.01 |
| City Hotel | Heavy Rain (>15mm) | 1667 | 22,1 | 35.39 | 109.44 | 117356.35 |
| Resort Hotel | Dry (0mm) | 30555 | 0.0 | 28.25 | 104.71 | 2198198.33 |
| Resort Hotel | Light Rain (0-5mm) | 5884 | 1,3 | 27,06 | 71.53 | 298406.82 |
| Resort Hotel | Moderate Rain (5-15mm) | 1771 | 8,5 | 30.77 | 63.66 | 77257.37 |
| Resort Hotel | Heavy Rain (>15mm) | 1098 | 23,5 | 24.59 | 64.59 | 53395.09 |

Surprisingly, heavy rain days have fewer cancellations than dry days (City: 35% vs 42%, Resort: 25% vs 28%). Maybe people who book knowing it'll rain are more committed to their trip? Either way, no reason to drop prices just because of rain.

Lead time against cancellation risk

| hotel | booking_window | bookings | cancel_rate_pct | avg_daily_rate | realized_revenue |
|--------------|--------------------------|----------|-----------------|----------------|------------------|
| City Hotel | Early Bird (3+ months) | 35048 | 55.87 | 102.7 | 1649171.95 |
| City Hotel | Advance (1-3 months) | 20662 | 40.08 | 108.02 | 1312004.04 |
| Resort Hotel | Early Bird (3+ months) | 15816 | 39.62 | 100.03 | 930841.00 |
| City Hotel | Short Notice (8-30 days) | 12393 | 31.21 | 114.65 | 961579.27 |
| City Hotel | Last Minute (0-7 days) | 10019 | 12.41 | 109.94 | 965668.44 |
| Resort Hotel | Advance (1-3 months) | 8645 | 32.75 | 106.14 | 600854.87 |
| Resort Hotel | Last Minute (0-7 days) | 8589 | 6.52 | 78.6 | 626846.17 |
| Resort Hotel | Short Notice (8-30 days) | 6258 | 22,05 | 100.53 | 468715.57 |

The earlier someone books, the more likely they cancel. City Hotel early bird bookings have 55.87% cancellation but last minute is only 12.41%. So people who book late actually show up.

Guest composition and weather preference

| hotel | guest_type | bookings | pct_of_hotel_bookings | cancel_rate | avg_rate | avg_stay_length | avg_temp_celsius | avg_precipitation_mm | avg_sunshine_hrs | total_revenue |
|--------------|---------------|----------|-----------------------|-------------|----------|-----------------|------------------|----------------------|------------------|---------------|
| City Hotel | Couple | 53111 | 67.98 | 45.27 | 101.16 | 3.0 | 17,8 | 1,4 | 10,6 | 3010655.49 |
| City Hotel | Solo Traveler | 14982 | 19.18 | 35.48 | 97.15 | 2,6 | 16,8 | 1,4 | 10,2 | 917600.00 |
| City Hotel | Family | 5121 | 6.56 | 34.64 | 154.26 | 3,4 | 18,4 | 1,1 | 11.0 | 508727.69 |
| City Hotel | Group | 4657 | 5.96 | 37.64 | 153.65 | 3,4 | 18,6 | 1,1 | 11,2 | 438637.25 |
| City Hotel | Other | 251 | 0.32 | 34.26 | 79.97 | 4.0 | 17,5 | 1,5 | 11,1 | 12803.27 |
| Resort Hotel | Couple | 27641 | 70.32 | 30.02 | 94.86 | 4,6 | 19,4 | 1,3 | 10,8 | 1802803.24 |
| Resort Hotel | Family | 3884 | 9.88 | 35.89 | 163.01 | 4,8 | 21,3 | 0.8 | 11,5 | 384620.47 |
| Resort Hotel | Solo Traveler | 6739 | 17.14 | 17.17 | 57.22 | 3,1 | 17.0 | 1,5 | 10.0 | 308512.71 |
| Resort Hotel | Group | 1043 | 2.65 | 18.22 | 156.36 | 5.0 | 21,6 | 0.5 | 11,6 | 131293.19 |
| Resort Hotel | Other | 1 | 0.0 | 0.0 | 28.0 | 10.0 | 13,9 | 0.0 | 8,7 | 28.0 |

Couples make up most bookings (around 70%). Families and groups come during warmer weather and spend more per booking. Solo travelers at Resort Hotel have really low cancellation (17%) — pretty reliable segment.

Service pricing estimate

| Estimate summary | | |
|------------------|--------------|-----------------------|
| Upfront cost | Monthly cost | Total 12 months cost |
| 0.00 USD | 1.59 USD | 19.08 USD |
| | | Includes upfront cost |

| Detailed Estimate | | | | |
|------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|--------------|--------------|
| Name | Group | Region | Upfront cost | Monthly cost |
| Amazon Simple Storage Service (S3) | - | US East (N. Virginia) | 0.00 USD | 0.02 USD |
| Status | - | | | |
| Description: | - | | | |
| Config summary | S3 Standard storage (1 GB per month) | | | |
| Name | Group | Region | Upfront cost | Monthly cost |
| AWS Glue | - | US East (N. Virginia) | 0.00 USD | 1.51 USD |
| Status | - | | | |
| Description: | - | | | |
| Config summary | Number of DPU's for Apache Spark job (10), Number of DPU's for Python Shell job (0.0625) | | | |
| Name | Group | Region | Upfront cost | Monthly cost |
| AWS Lambda | - | US East (N. Virginia) | 0.00 USD | 0.00 USD |
| Status | - | | | |
| Description: | - | | | |
| Config summary | Architecture (x86), Architecture (x86), Invoke Mode (Buffered), Amount of ephemeral storage allocated (512 MB), Number of requests (4 per month) | | | |
| Name | Group | Region | Upfront cost | Monthly cost |
| Amazon Athena | - | US East (N. Virginia) | 0.00 USD | 0.06 USD |
| Status | - | | | |
| Description: | - | | | |
| Config summary | Total number of queries (20 per day), Amount of data scanned per query (20 MB) | | | |
| Name | Group | Region | Upfront cost | Monthly cost |
| Amazon EventBridge | - | US East (N. Virginia) | 0.00 USD | 0.00 USD |
| Status | - | | | |
| Description: | - | | | |
| Config summary | Size of the payload (1 KB), Number of events (4 per month) | | | |

Suggestions

Don't discount rooms when rain is forecasted — the data shows these bookings actually cancel less (7% lower). These guests probably planned for it.

Wind hurts Resort Hotel prices a lot (drops to \$64 from \$100). Maybe send guests early updates when strong wind is expected and adjust rates accordingly.

Early bookings cancel way more (55%), so it makes sense to require higher deposits for those. Last minute bookings are reliable so keep those flexible.

Families and groups pay more (~€154-163 vs €95 average) and prefer warm sunny weather. Could push marketing to them when forecasts look good.

Thank You