Wildfire Footprints: Cloud-Based Forecasting with NASA MODIS Data

- **Leveraging EC2, RDS, and Jupyter on AWS (2003, 2013, 2023)**
- Team Members: Ritu Patel & Dhwanil Mori

Project Scope & Motivation

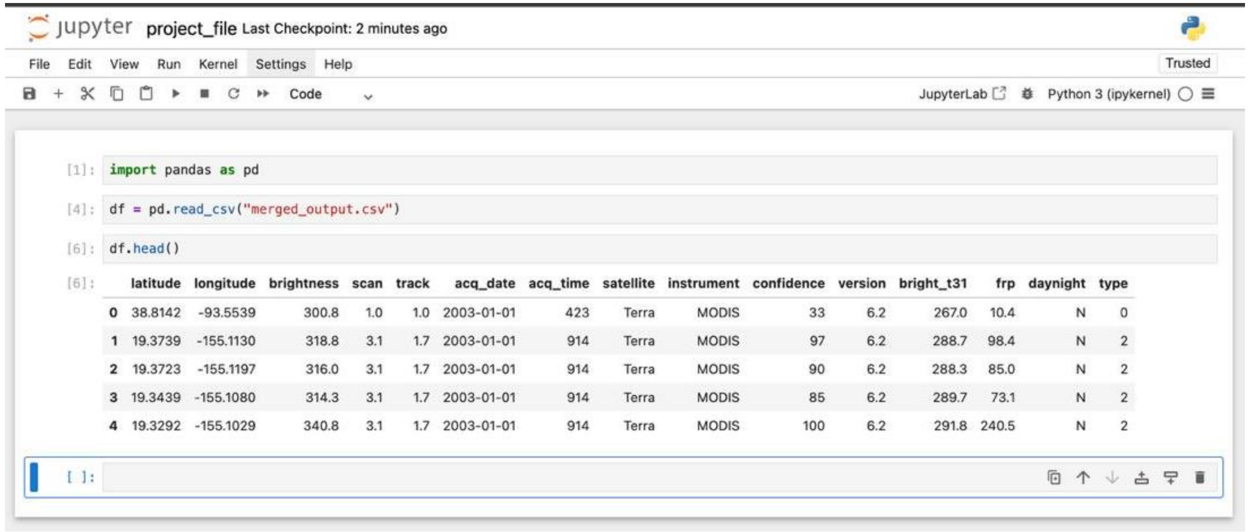
- Analyze wildfire trends across two decades (brightness and FRP).
- Build predictive models for both intensity (**FRP**) and heat (**Brightness**).
- Motivation: Improve wildfire preparedness and risk assessment.

Cloud Architecture

1	Upload CSVs to EC2	AWS	Jupyter Notebook Server
2	Create RDS MySQL DB	AWS	Stores wildfire data
3	Load Data into RDS	AWS (via Jupyter)	Python Scripts
4	Analyze and Model	AWS (Jupyter)	VAR + LSTM
5	Cloud-Native Workflow	✅	No local machine needed

Data Sources & Key Features

- **Data Source:** NASA MODIS (Terra and Aqua Satellites)
- **Key Variables:**
- Fire Radiative Power (FRP)
- Brightness
- Confidence Score
- Latitude/Longitude
- **Years:** 2003, 2013, 2023
- **Volume:** ~100,000 records per year



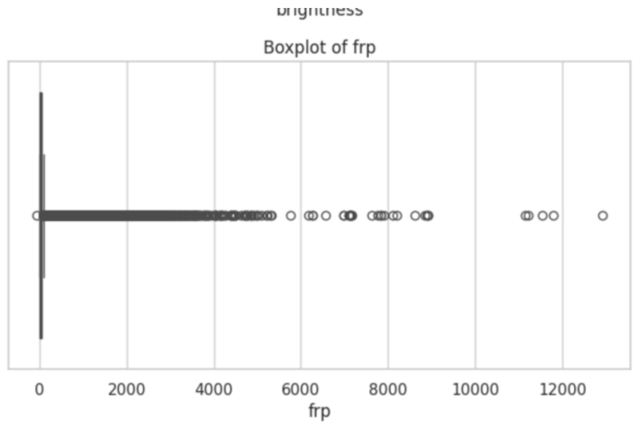

```
[1]: import pandas as pd
[4]: df = pd.read_csv("merged_output.csv")
[6]: df.head()
```

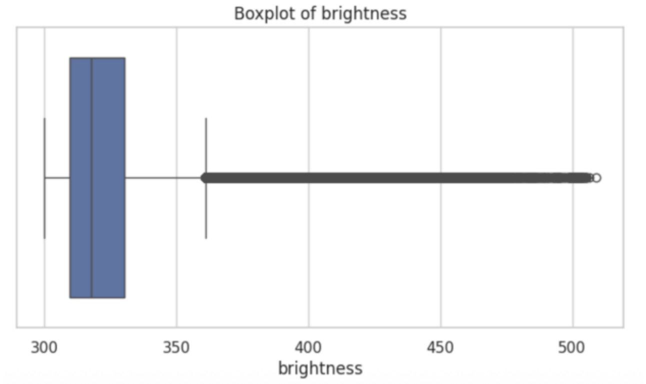
[6]:

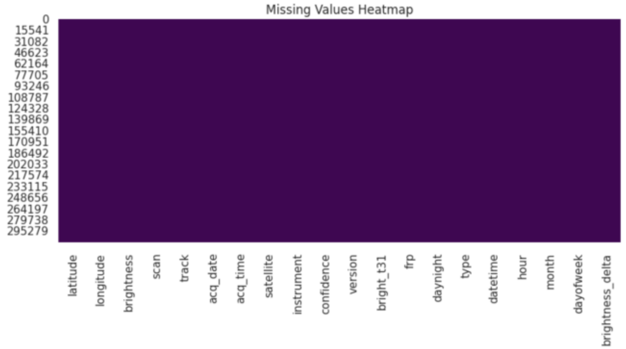
	latitude	longitude	brightness	scan	track	acq_date	acq_time	satellite	instrument	confidence	version	bright_t31	frp	daynight	type
0	38.8142	-93.5539	300.8	1.0	1.0	2003-01-01	423	Terra	MODIS	33	6.2	267.0	10.4	N	0
1	19.3739	-155.1130	318.8	3.1	1.7	2003-01-01	914	Terra	MODIS	97	6.2	288.7	98.4	N	2
2	19.3723	-155.1197	316.0	3.1	1.7	2003-01-01	914	Terra	MODIS	90	6.2	288.3	85.0	N	2
3	19.3439	-155.1080	314.3	3.1	1.7	2003-01-01	914	Terra	MODIS	85	6.2	289.7	73.1	N	2
4	19.3292	-155.1029	340.8	3.1	1.7	2003-01-01	914	Terra	MODIS	100	6.2	291.8	240.5	N	2

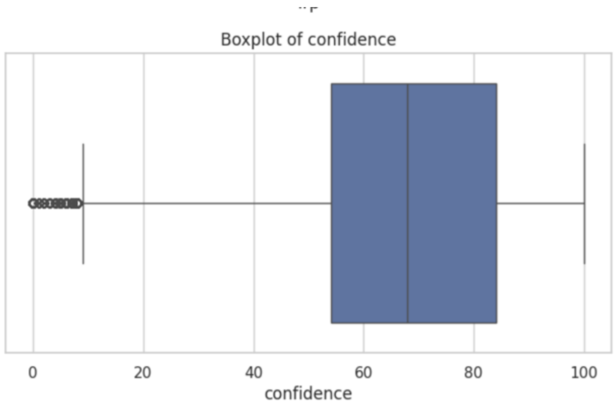
```
[ ]:
```

Feature Engineering

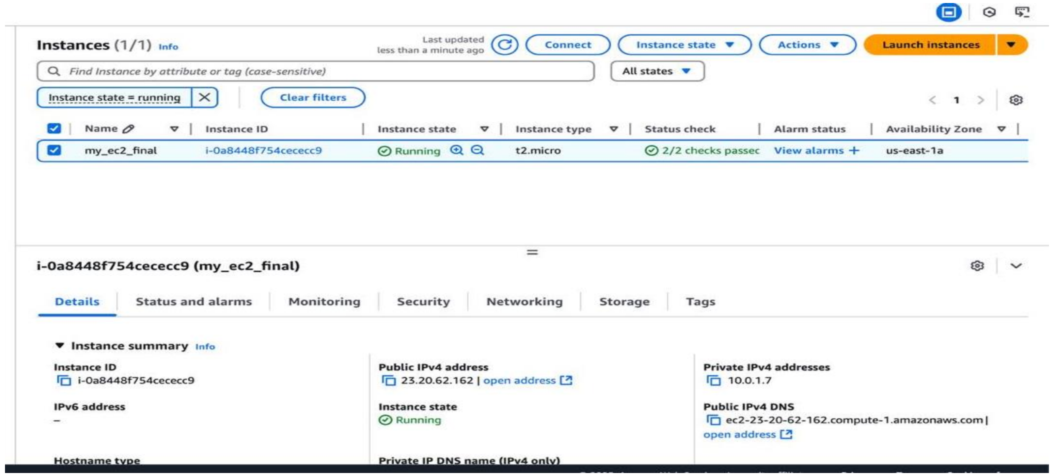




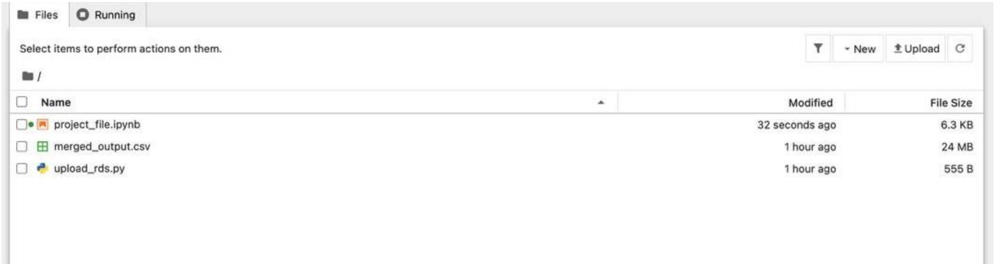


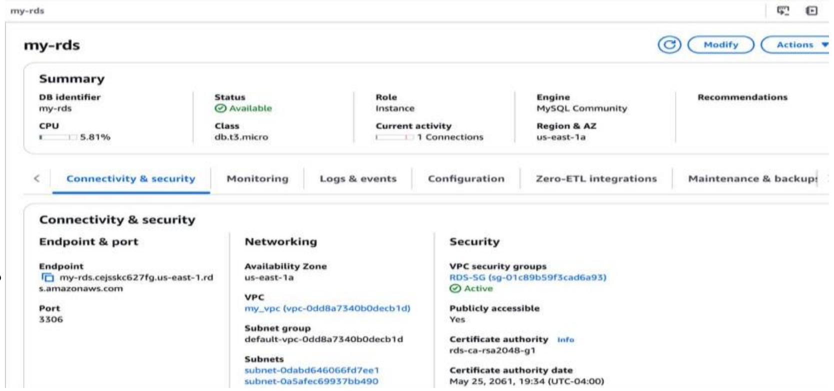


Data Processing Pipeline



- **Upload** raw CSVs to EC2.
- **Store** structured data in RDS MySQL.
- **Connect** Jupyter to RDS for EDA and preprocessing.
- **Feature Engineering**:
- Temporal Features (Month, Year)
- Categorized Fire Intensity
- Brightness Normalization
- **Prepare** input sequences for LSTM and VAR modeling.





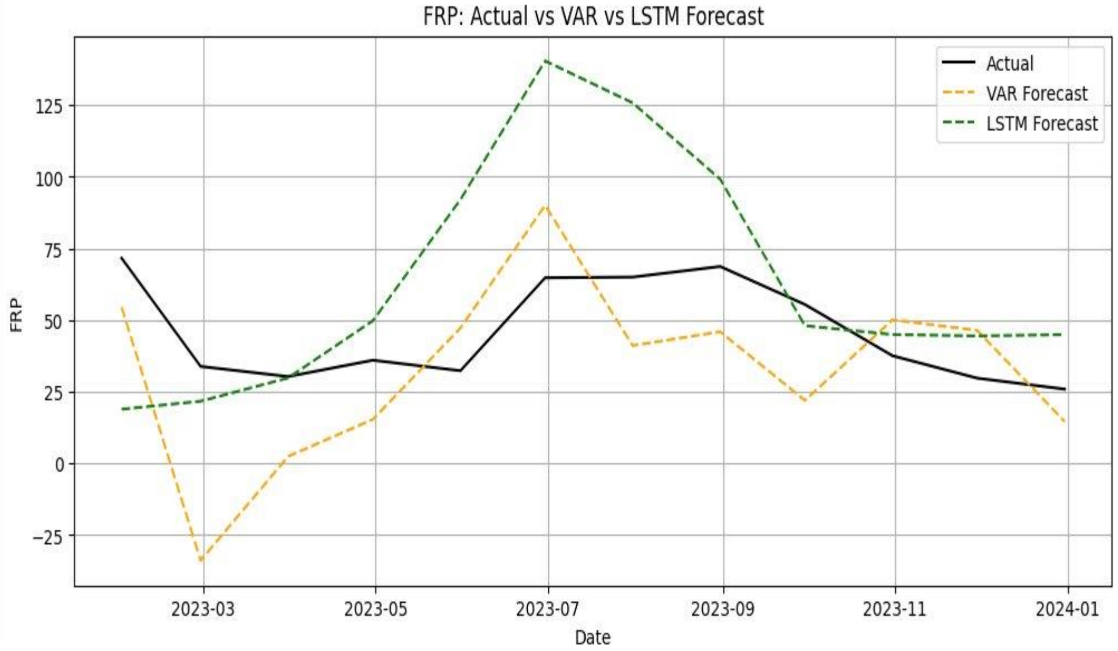
Forecasting Approaches: VAR and LSTM)

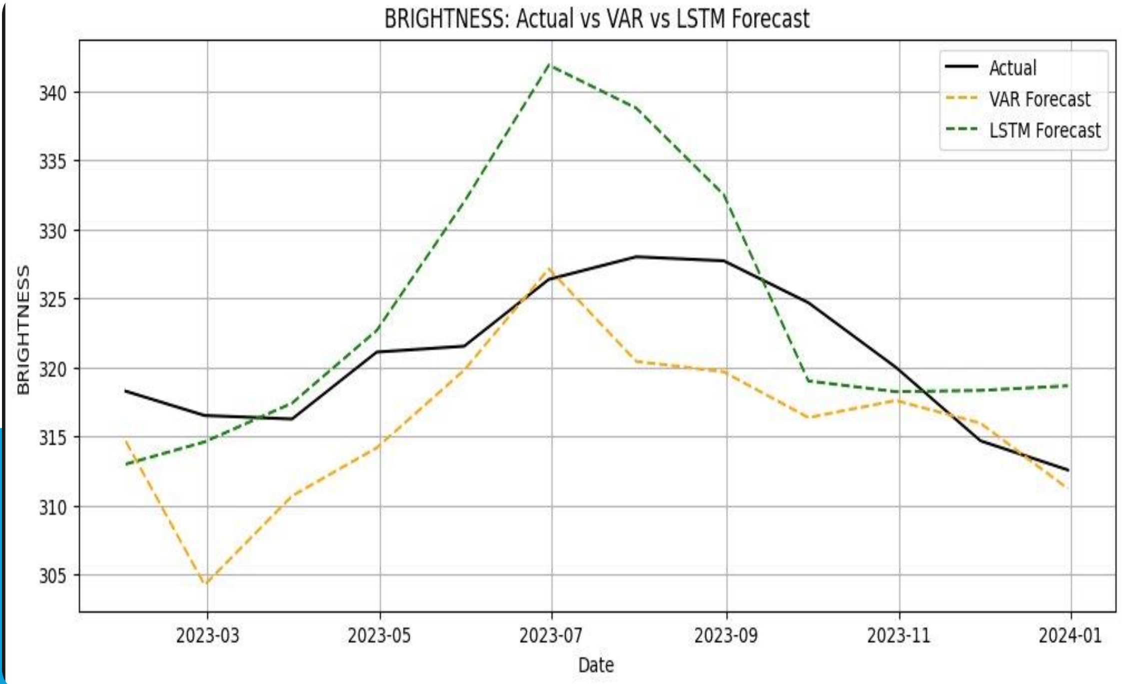
- **VAR (Vector AutoRegression)**:
  - Classical time-series model assuming linear relationships.
  - Captures multi-variable dependencies.
- **LSTM (Long Short-Term Memory)**:
  - Deep learning model for sequential data.
  - Captures complex, nonlinear patterns and long-term dependencies.

Model	Variable	RMSE	MAE	MAPE	R^2
VAR	Brightness	7.02	5.47	11.72%	0.96
	FRP	30.12	24.86	21.10%	0.89
LSTM	Brightness	7.13	5.72	13.49%	0.94
	FRP	38.43	29.53	25.61%	0.86

Model Results & Metrics

- **VAR** outperformed LSTM in both FRP and Brightness forecasting
- LSTM showed higher error and lower R^2 due to FRP overfitting
- Brightness predictions were more stable across both models





COMPARATIVE FORECAST ANALYSIS

- VAR:
- Performs well on smoother signals like Brightness
- Generates consistent forecasts with low volatility

- LSTM:
- Captures complex seasonal patterns
- Overfits to FRP peaks, producing exaggerated forecasts

- While LSTM captures deeper temporal dependencies, VAR provides more reliable and interpretable outputs for operational use cases.

Challenges Faced

- **Data Quality**:
  Handling missing values and low-confidence wildfire detections.
- **Cloud Setup**:
  Configuring secure connections between EC2, RDS, and Jupyter.
- **Forecasting Volatility**:
  Fire Radiative Power (FRP) was highly unstable, making predictions difficult.

Future Work

- **Migrate workflows to Amazon SageMaker** for scalable model training and deployment.
- **Automate real-time data ingestion** using AWS Lambda and Amazon Kinesis.
- **Optimize storage** by archiving older wildfire data to Amazon S3 Glacier.
- **Deploy forecasting models** as APIs using AWS Elastic Beanstalk or API Gateway.
- **Enhance security and monitoring** with IAM roles, Secrets Manager, and CloudWatch.