

# IT Certifications Pipeline

*Student:* **Zakaria Ait Hssain**

*Supervisor:* **Pr. M. Lazaar**

End of Year Project

June 2, 2025

# Contents

---

1. Introduction
2. Data Collection (Scraping)
3. Data Cleaning
4. Machine Learning Enrichment
5. Data Analysis
6. Streamlit UI
7. Conclusion

# Introduction

---

# Problem & Objectives

---

## Problem

- IT certifications are critical in hiring and skill demonstration.
- No unified source or pattern across providers.
- Thus: difficulty to compare and choose between certifications.

## Objectives

Create a unified and accessible certification dataset to enable data-driven analysis and choice.

# Data Flow CCM (Actor/Flow)

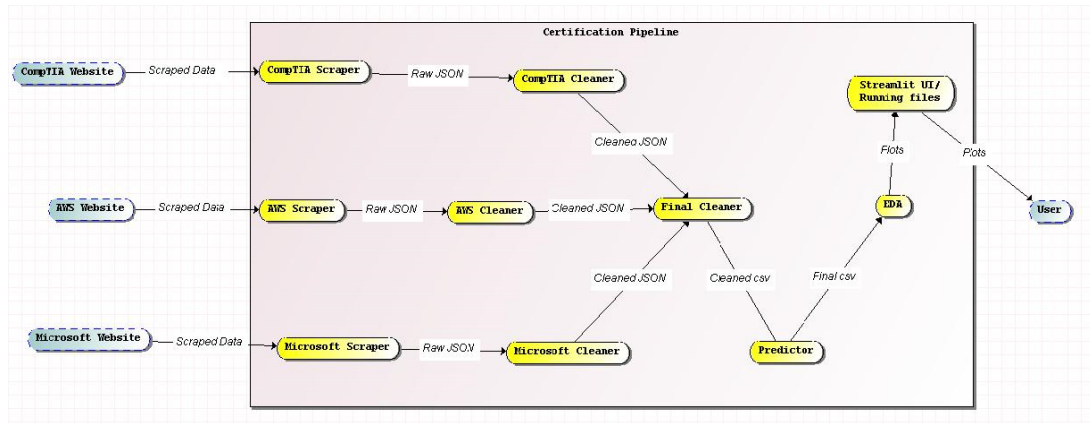


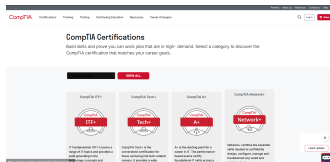
Figure: Communication Conceptual Model

## **Data Collection (Scraping)**

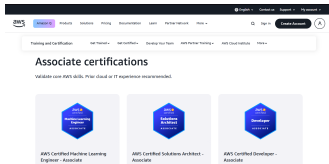
---

# Certifications Boxes Page

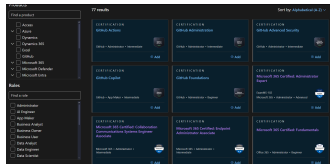
## CompTIA



## AWS



## Microsoft



# Certification Details Page

## CompTIA

### Exam Details

Exam Codes	Exam FCO-U81
Launch Date	September 4, 2018
Exam Details	The CompTIA IT Fundamentals (ITF+) exam focuses on the knowledge and skills required to identify and explain the basics of computing, IT infrastructure, software development and database use.
Number of Questions	Maximum of 75 questions per exam
Type of Questions	Multiple choice
Length of Test	60 Minutes
Passing Score	650 (on a scale of 800)
Recommended Experience	No prior experience necessary
Languages	English
Retirement	Usually three years after launch
Testing Provider	Pearson VUE <ul style="list-style-type: none"><li>• Testing Centers</li><li>• Online Testing</li></ul>
Price	<a href="#">View price in cart</a> (see all pricing)

#### Get Practice Questions and Exam Objectives

First NameLast Name

EmailJob Description

Select Exam

Country

☐ I'm interested in receiving:  
☐ Exam Objectives☐ Practice Questions

Training Status

I plan on taking the exam...

☐ I agree to the [Terms of Use & Privacy statement](#).

SUBMIT

Label options

CD

100%

OVERVIEW

EXAM DETAILS

PREPARATION

RENEWAL

TRY OUR SIMULATED

BUY NOW

## AWS

aws

Amazon Q

Products

Solutions

Pricing

Documentation

Learn

Partner network

More

English

Contact us

Support

My account

Q

Sign in

Create Account

A

Training and Certification

Get Trained

Get Certified

Develop Your Team

AWS Partner Training

Education Programs

More

### Exam overview

AWS Certified Machine Learning Engineer - Associate

### Category

Associate

### Exam duration

150 minutes

### Exam format

65 questions

### Cost

150 USD. Visit [Exam pricing](#) for additional cost information, including foreign exchange rates.



# Certification Details Page

## Microsoft First Layout

Microsoft Certified: Security Operations Analyst Associate

Option 3: Prerequisite certification

### Certification exams

Exam SC-900: Microsoft Cybersecurity Architect

Languages: English, Japanese, Chinese (Simplified), Korean, German, French, Spanish, Portuguese (Brazil), Chinese (Traditional), Italian

Refinement date: none

This exam measures your ability to accomplish the following technical task: design solutions that align with security best practices and priorities; design security operations, identity, and compliance capabilities; design security solutions for infrastructure and design security solutions for applications and data.

[Schedule exam >](#)

ⓘ We strongly recommend that you register for an exam with a personal MCA account. If you register with an organizational (work/school) AD account, your exam records will be lost if you leave your organization and they will be unrecoverable.

Take a free practice assessment

Test your skills with practice questions to help you prepare for the exam. Learn more about practice assessments.

United States

\$165 USD\*

Price based on the country or region in which the exam is proctored.

## Microsoft Second Layout

### At a glance

Level  
Intermediate

Product  
Windows

Role  
Business User

### Jump to

- [Prepare for the exam](#)
- [Take the exam](#)
- [Certification resources](#)

[Show less](#)

### Overview

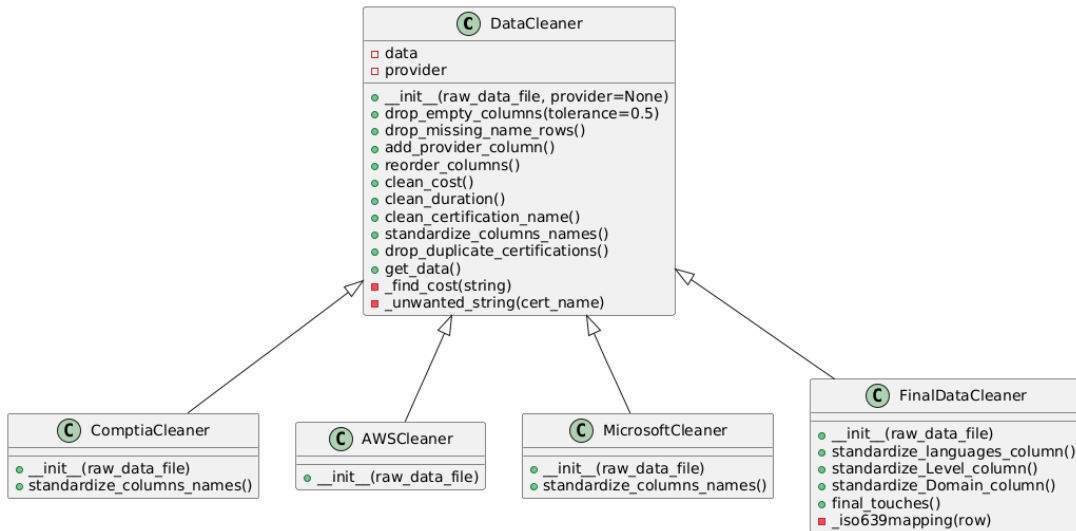
The Microsoft Certified Educator (MCE) certification validates that you have the global educator technology literacy competencies needed to provide a rich, custom learning experience for students. This certification is ideal for:

- Educators-in-training
- Faculty of teacher training colleges
- In-service educators

# Data Cleaning

---

# Cleaning Classes Diagram



# Machine Learning Enrichment

---

# Machine Learning Enrichment

---

- Fields predicted: Level, Domain, Cost, and Duration
- Models used: Decision Trees and Random Forests (for duration)
- Performed: One-Hot encoding, Feature Selection, and Class weighting
- Dimensionality Reduction (Specifically PCA) was not necessary

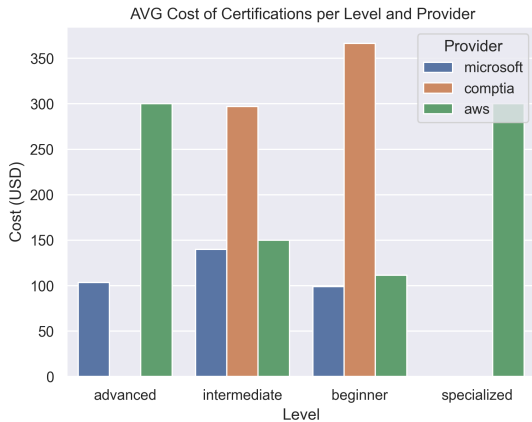
**Table:** Best cross-validation scores for each predicted field

Target Field	Used Model	Best CV Score
Level (classifier)	Decision Tree	0.869
Domain (classifier)	Decision Tree	0.459
Cost (regressor)	Decision Tree	0.93
Duration (regressor)	Random Forest	0.837

# Data Analysis

---

# Key Insights and Analysis



- CompTIA certifications presents the highest average costs, surprisingly at the Beginner level.
- AWS shows a clear cost progression by level.
- Microsoft maintains relatively stable and moderate pricing.

Figure: Average Cost per Level and Provider

# Key Insights and Analysis

---

## General Insights:

- Cost range: \$100 – \$392
- Duration modality: 60 & 100 min
- Dominant level: Intermediate
- Most common domains: Productivity, Business Apps & Data.
- Most common languages: English, Japanese & Spanish

*More plots and analysis are present in the project report.*



## Streamlit UI

---

# Key Features

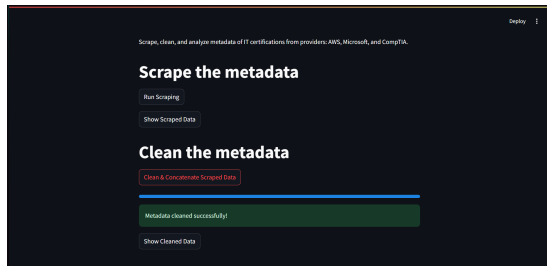
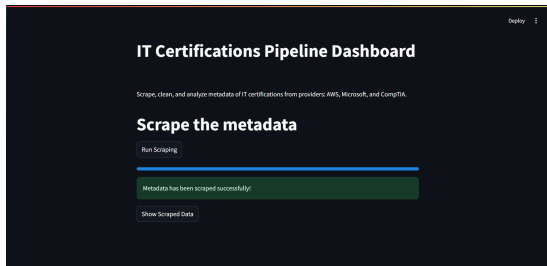
---

## Key features

- Run full pipeline step by step.
- Track progress.
- View datasets and plots.
- Download datasets.

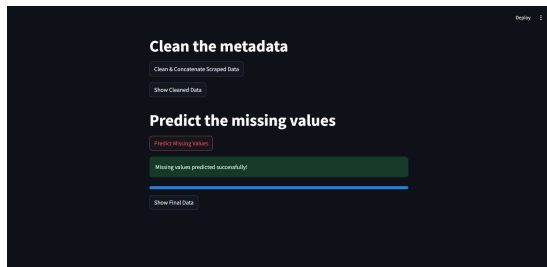
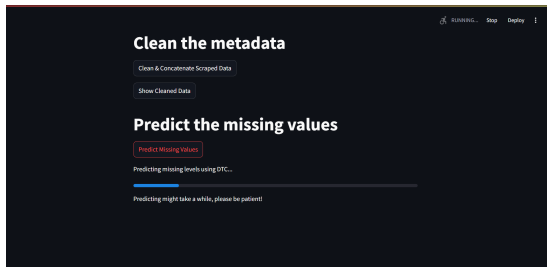
# UI Overview

---



# UI Overview

---



# UI Overview

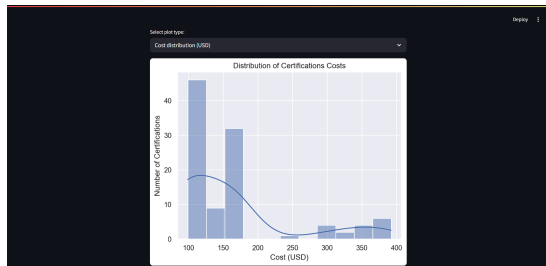
Deploy

Show Final Data

Final Data:

	Certification	Provider	Domain	Level	Languages
0	365 certified administrator	microsoft	productivity	advanced	english, chinese (simplified)
1	365 certified collaboration communic	microsoft	productivity	intermediate	english, chinese (simplified)
2	365 certified endpoint administrator	microsoft	productivity	intermediate	english, chinese (simplified)
3	365 certified fundamentals	microsoft	productivity	beginner	english, japanese, spanish
4	365 certified teams administrator	microsoft	productivity	intermediate	english, japanese, chinese
5	ai+	compta	business apps	beginner	english, german, japanese
6	comp+	compta	data & other	beginner	not specified
7	certified advanced networking	aws	data & other	beginner	not specified
8	certified ai practitioner	aws	it/business	beginner	english, japanese, korean
9	certified azure administrator	microsoft	cloud	intermediate	english, chinese (simplified)

Explore



## Conclusion

---

# Conclusion

---

## This project offers

- Automated, trackable pipeline
- Unified, rich certification dataset
- Valuable tool for certification seekers

## Possible project extensions

- Add more providers (e.g., Cisco, Google Cloud, Oracle)
- Job market integration (salary, demand)
- ...

# Thank you!

zakaria\_aithssain@um5.ac.ma

<https://github.com/zakariaaithssain/it-certifications-pipeline>