****

**Table of Contents:**

[**1. Introduction: 3**](#_fwjb3ond9woi)

[**2. Business Questions: 3**](#_5u8sg85lpfgx)

[**3. Database Design: 4**](#_ep6rv9gve0dv)

[3.1. Entity Relationship Diagram(ERD): 4](#_m268eaiajcuc)

[3.2. The Schema: 4](#_er7iz0r1fi6m)

[**4. Database Creation: 5**](#_rtin6n1ndu4t)

[4.1 Database Tables Creation: 5](#_s0ee3le8w126)

[4.2 Creation of Table View 7](#_hl0jus35dhti)

[4.3 Data Insertion 8](#_tfizkaw61edd)

[4.4 Data Selection 10](#_elotqka658sq)

[**5. Exploratory Data Analysis(EDA): 13**](#_vf81s5b9npqu)

[**6. Pivot Tables: 14**](#_jkjylvsfo18o)

[6.1 Analysis and Reports Overall: 14](#_820pewvpator)

[6.1.1. Count of overall projects for each trainer 14](#_6nbhjge0ym77)

[6.1.2. Records of the hourly rate and satisfaction rate per trainer and projects throughout the whole years 15](#_fdrhh6rkmht0)

[6.1.3. Reports per subject 15](#_j94ai7e8an0k)

[6.1.4. Data per Gender 16](#_iawq936ddden)

[6.1.5. Number of hours vs the average hourly rate 16](#_cdmz79we1oki)

[6.1.6. Project data over years 17](#_axskxjqjvxgg)

[6.1.7. Organizations and Projects 17](#_akbwunr27ez4)

[6.1.8. View over experience level 18](#_6rjd66cy3rew)

[6.1.9. Last active year and the hourly rate corresponding to this year for each trainer 18](#_jgmjoa6wk0mk)

[6.2 Analysis and Reports for the current year: 19](#_7t87peapvsg5)

[6.2.1.Trainer and project\_rate per project for the current year 19](#_f163q9gpjtjb)

[6.2.2.Average hourly rate for the current year per trainer 19](#_ia9vq43npsos)

[6.2.3.Average hourly rate for the current year per trainer 20](#_2vsm0djnwo69)

[6.2.4.Average hourly rate vs years of experience for the current year 20](#_q4c0rvd2lv)

[6.2.5.Project data for the current year 21](#_bjij7u456f9q)

[6.2.6.Details of projects and trainers for the current year 21](#_hzsnw238rk0a)

[6.2.7.Trainer and Organization 21](#_i29krsuycgvm)

[6.2.8.View over experience level for the current year 22](#_ndleo0j358pq)

[6.3. A report of the trainer’s raw data 22](#_ozshn7ferh12)

[**7. Analysis and Graphs: 23**](#_krd77qj3001x)

[**8. Recommendations: 24**](#_7a6kaifgnz9g)

# **1. Introduction:**

The **ICEALEX** team asked to do a project to manage their data and optimize trainers’ allocation across **ICEALEX** programs. This project aims to enhance the effectiveness of **ICEALEX'**s training programs by focusing on two key objectives:

* **Analyzing Trainer Satisfaction Rates:** By measuring and improving trainer satisfaction, we aim to boost the quality of education provided, ensuring that satisfied trainers create more engaging and impactful learning experiences for trainees.
* **Creating Comprehensive Trainer Portfolios:** Developing detailed portfolios for each trainer will capture their expertise, qualifications, performance metrics, and program history. This will enable data-driven decisions for optimal trainer deployment, fostering continuous improvement and excellence in training delivery.

**Some issues were encountered throughout the project**:

Inconsistent Data Entries: Several trainer records had missing or inconsistent data, making it difficult to assess performance and qualifications accurately.

**The tools and methodologies used in the project:**

* SQL Server is utilized to create and manage the database, ensuring efficient storage and handling of trainer data and relationships between different data tables.
* Python: Utilized to do Exploratory Data Analysis (EDA) to explore the data, identify patterns, and address inconsistencies.
* Excel: Used to construct pivot tables and connect with the SQL database, facilitating a detailed and dynamic examination of the data.
* Power BI: Applied to analyze the data and create interactive visualizations. Power BI allowed us to present the data clearly and engagingly, helping stakeholders to easily understand the insights and trends.

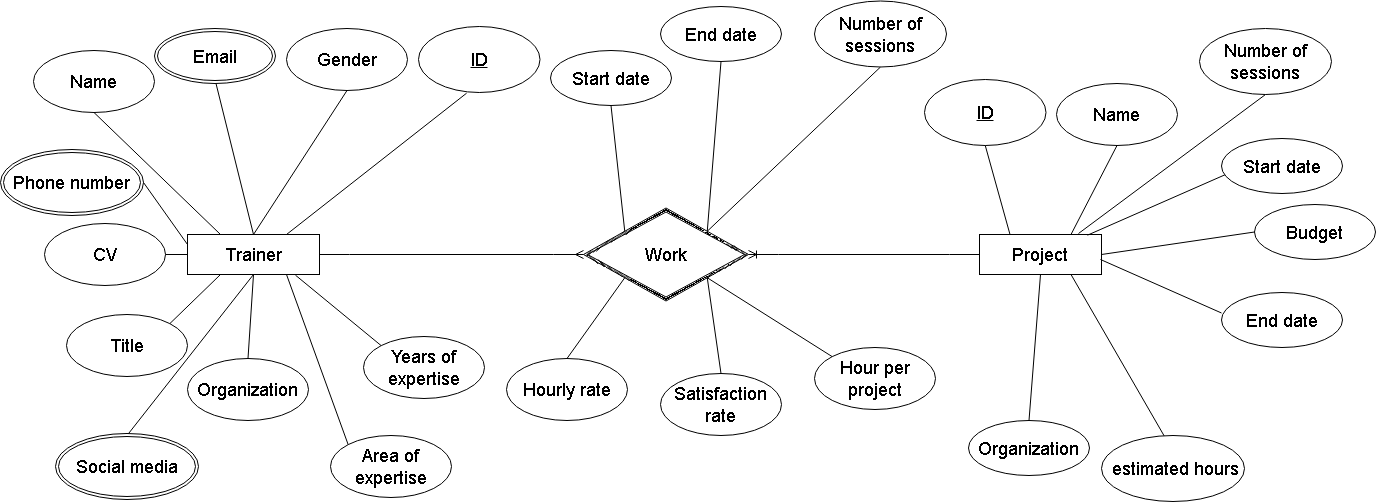
Through addressing these issues and implementing our objectives, we strive to elevate both trainer and trainee success, driving the overall mission of ICEALEX forward.

# **2. Business Questions:**

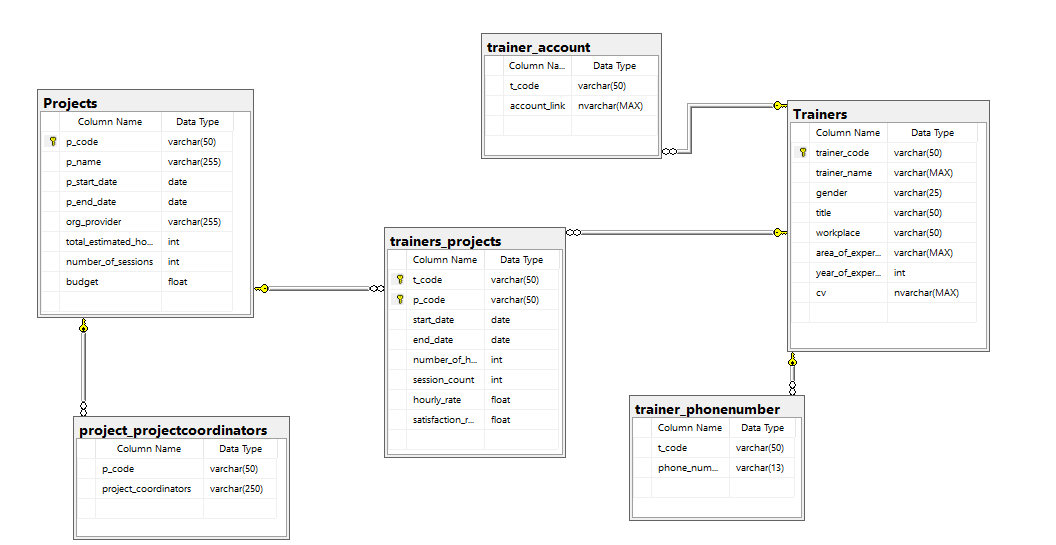
1. The ICEALEX community has several trainers. Each trainer has a name, gender, ID, CV, title, organization, years of experience, area of expertise, satisfaction rate, hourly rate, hour per project, multiple emails, phone numbers, and social media accounts.
2. The community has many projects. Each project has a name, ID, number of sessions, start date, end date, budget, organization, and estimated hours.
3. Each trainer may work on many projects
4. Each project must done by one or more trainer

# 3. Database Design:

## **3.1. Entity Relationship Diagram(ERD):**



## **3.2. The Schema:**



# **4. Database Creation:**

In this step, we used SQL Server to create, insert, and select the trainer’s data.

**The tables created are:**

1. Trainers table: A dimension table that contains the trainers' data.
2. trainer\_phonenumber table: A dimension table containing multiple trainers’ phone number values.
3. trainer\_account: A dimension table that contains multiple values of trainers’ email addresses and social media accounts.
4. Projects: A dimension table that contains IceAlex’s projects.
5. Project\_projectcoordinators: A dimension table containing the coordinators' names, assigned for the projects.
6. trainers\_projects table: A fact table containing both trainers’ data and the projects they worked on.

| **Table name** | **Key** | **Type** |
| --- | --- | --- |
| Trainers | trainer\_code | Primary key |
| trainer\_phonenumber | t\_code,phone\_number | Composite key |
| trainer\_account | t\_code,account\_link | Composite key |
| Projects | p\_code | Primary key |
| project\_projectcoordinators | p\_code,project\_coordinator | Composite key |
| trainers\_projects | t\_code,p\_code | Composite key |

## 4.1 Database Tables Creation:

**Create trainer portfolio database:**

CREATE DATABASE trainer\_portfolio;

USE trainer\_portfolio;

**Create Trainers table:**

CREATE TABLE Trainers(

trainer\_code varchar(50) PRIMARY KEY,

trainer\_name varchar(max) NOT NULL,

gender varchar(25) NOT NULL,

title varchar(50),

workplace varchar(50),

area\_of\_expertise varchar(max),

year\_of\_expertise int,

cv nvarchar(max) )

**Create trainer-accounts table:**

CREATE TABLE trainer\_account(

t\_code varchar(50) FOREIGN KEY REFERENCES Trainers(trainer\_code),

account\_link nvarchar(max)

)

**Create trainer-phone number table:**

CREATE TABLE trainer\_phonenumber(

t\_code varchar(50) FOREIGN KEY REFERENCES Trainers(trainer\_code),

phone\_number varchar(13),

)

**Create projects table:**

CREATE TABLE Projects(

p\_code varchar(50) PRIMARY KEY,

p\_name varchar(255) NOT NULL,

p\_start\_date Date,

p\_end\_date Date,

org\_provider varchar(255),

total\_estimated\_hours int,

number\_of\_sessions int,

budget Float

)

**Create a project coordinator table:**

CREATE TABLE project\_projectcoordinators(

p\_code varchar(50) FOREIGN KEY REFERENCES Projects(p\_code),

project\_coordinators varchar(250)

)

**Create trainers\_projects table:**

CREATE TABLE trainers\_projects(

t\_code varchar(50) FOREIGN KEY REFERENCES Trainers(trainer\_code),

p\_code varchar(50) FOREIGN KEY REFERENCES Projects(p\_code),

start\_date Date,

end\_date Date,

number\_of\_hours\_per\_project int NOT NULL,

session\_count int,

hourly\_rate Float,

satisfaction\_rate Float,

PRIMARY KEY(t\_code,p\_code)

)

## 4.2 Creation of Table View

**Create a view to get the email address only:**

CREATE VIEW v\_email AS

SELECT t.trainer\_code, t.trainer\_name, tc.account\_link

FROM Trainers AS t

LEFT JOIN trainer\_account AS tc

ON t.trainer\_code = tc.t\_code

WHERE tc.account\_link LIKE '%@%';

**Create a view to get the social media only:**

CREATE VIEW v\_socialmedia AS

SELECT t.trainer\_code, t.trainer\_name, tc.account\_link

FROM Trainers AS t

LEFT JOIN trainer\_account AS tc

ON t.trainer\_code = tc.t\_code

WHERE tc.account\_link LIKE '%www.%';

**Create a view for the current year:**

CREATE VIEW current\_year AS

SELECT \*

FROM trainers\_projects

WHERE year(start\_date) = (SELECT MAX(year(start\_date)) FROM trainers\_projects);

**Create a view to have both trainers and project names together in the trainer view:**

CREATE VIEW detailed\_trainers\_projects AS

SELECT tp.t\_code,t.trainer\_name,tp.p\_code, p.p\_name,t.area\_of\_expertise,tp.start\_date,tp.end\_date,

tp.number\_of\_hours\_per\_project,tp.session\_count,tp.hourly\_rate,tp.satisfaction\_rate

FROM trainers\_projects AS tp

LEFT JOIN Trainers AS t

ON t.trainer\_code = tp.t\_code

LEFT JOIN Projects AS p

ON p.p\_code = tp.p\_code;

## 4.3 Data Insertion

**Inserting data to Trainers table:**

INSERT INTO Trainers(trainer\_code,trainer\_name,gender,title,workplace,area\_of\_expertise,year\_of\_expertise,cv)

VALUES

('Trainer01','AAAAAA','Female','Business Developer','Na','Business Development',4,'https://www.google.com/drive/'),

('Trainer02','BBBBBB','Male','Business Consultant','Na','(Business Model) Innovation Consultancy',4,'<https://www.google.com/drive/>'),

|

|

|

**Insert data into trainer\_account:**

INSERT INTO trainer\_account(t\_code,account\_link)

VALUES

('Trainer01','trainer@gmail.com'),

('Trainer02','trainer@gmail.com'),

('Trainer03','trainer@gmail.com'),

('Trainer04','trainer@gmail.com'),

('Trainer05','trainer@gmail.com'),

('Trainer06','trainer@gmail.com'),

('Trainer07','trainer@gmail.com'),

|

|

|

**Insert data into the phone number table:**

INSERT INTO trainer\_phonenumber(t\_code, phone\_number)

VALUES

('Trainer01','00000000000'),

('Trainer02','00000000000'),

('Trainer03','00000000000'),

('Trainer04','00000000000'),

('Trainer05','00000000000'),

('Trainer06','00000000000'),

('Trainer07','00000000000'),

('Trainer08','00000000000'),

|

|

|

**Insert data into the project table:**

INSERT INTO Projects(p\_code,p\_name,p\_start\_date,p\_end\_date,org\_provider,total\_estimated\_hours,number\_of\_sessions,budget)

VALUES

('SOA23','Startups of Alex','2023-01-01','2023-06-01','Na',40, 10, 5000),

('SOA22','Startups of Alex','2022-01-01','2022-06-01','Na',30, 10, 3000),

('SOA24','Startups of Alex','2024-01-01','2024-06-01','Na',40, 10, 6000),

('TIEC24','TIEC Pre-incubation','2024-01-01','2024-06-01','Na',40, 10, 5000),

('TIEC23','TIEC Pre-incubation','2023-01-01','2023-06-01','Na',40, 10, 5000),

('TIEC21','TIEC Pre-incubation','2021-01-01','2021-06-01','Na',40, 10, 1000),

('HR23','Heya Raeda','2023-01-01','2023-04-01','Na',90, 45, 10000),

|

|

|

**Insert data to project\_project\_projectcoordinators**

INSERT INTO project\_projectcoordinators(p\_code,project\_coordinators)

VALUES

('SOA23','AAA'),

('SOA22','AAA'),

('SOA24','BBB'),

('TIEC24','BBB'),

('TIEC23','CCC'),

('TIEC21','HHHH'),

('HR23','SSSS'),

('HR22','ZZZ'),

('HR21','ZZZ'),

|

|

|

**Insert data into trainers\_project table:**

INSERT INTO trainers\_projects(t\_code, p\_code,start\_date,end\_date,number\_of\_hours\_per\_project,session\_count,hourly\_rate, satisfaction\_rate)

VALUES

('Trainer01','SOA23','2023-01-01','2023-02-01',32,8,400,0.81),

('Trainer01','TIEC23','2023-02-01','2023-03-01',9,3,100,0.81),

('Trainer01','HR23','2023-03-01','2023-04-01',48,12,400,0.81),

('Trainer02','SOA22','2022-01-01','2022-02-01',32,8,300,0.52),

('Trainer02','TIEC24','2024-02-01','2024-03-01',9,3,600,0.2152),

('Trainer02','HR22','2022-03-01','2022-04-01',48,12,500,0.30),

|

|

|

## 4.4 Data Selection

**Show the Trainers data**

SELECT\*

FROM Trainers;

**Show the trainer\_account data:**

SELECT\*

FROM trainer\_account;

**Select the trainer data and trainer\_account:**

SELECT t.trainer\_code, t.trainer\_name, tc.account\_link

FROM Trainers AS t

LEFT JOIN trainer\_account AS tc

ON t.trainer\_code = tc.t\_code;

**Selection of the v\_email:**

SELECT \*

FROM v\_email

ORDER BY trainer\_code;

**Selection of the v\_social media:**

SELECT \*

FROM v\_socialmedia

ORDER BY trainer\_code;

**Show data from the phone number table:**

SELECT \*

FROM trainer\_phonenumber;

**Join trainers and phone number tables:**

SELECT t.trainer\_code, t.trainer\_name, tpn.phone\_number

FROM Trainers AS t

LEFT JOIN trainer\_phonenumber AS tpn

ON t.trainer\_code = tpn.t\_code

ORDER BY t.trainer\_code;

**Show the projects table:**

SELECT \*

FROM Projects;

**Show the project\_projectcoordinators:**

SELECT \*

FROM project\_projectcoordinators;

**Join the data in the project\_projectcoordinators with the Projects:**

SELECT p.p\_code, p.p\_name, pc.project\_coordinators

FROM Projects AS p

LEFT JOIN project\_projectcoordinators AS pc

ON p.p\_code = pc.p\_code

ORDER BY p.p\_code;

**Show trainers\_project table:**

SELECT \*

FROM trainers\_projects;

**Show data for the current year:**

SELECT \*

FROM current\_year;

**Show data for the detailed\_trainers\_projects:**

SELECT \*

FROM detailed\_trainers\_projects;

**Truncate database tables**

**Trainers Table**

TRUNCATE TABLE Trainers;

**trainer\_account**

TRUNCATE TABLE trainer\_account;

**trainer\_phonenumber**

TRUNCATE TABLE trainer\_phonenumber;

**Projects table**

TRUNCATE TABLE Projects;

**project\_projectcoordinators table**

TRUNCATE TABLE project\_projectcoordinators;

**trainers\_projects table**

TRUNCATE TABLE trainers\_projects;

**To update a value**

UPDATE Trainers

SET year\_of\_expertise = 4

WHERE trainer\_name = "AAAAA" ;

**To add new records**

INSERT INTO Trainers(trainer\_code,trainer\_name,gender,title,workplace,area\_of\_expertise,year\_of\_expertise,cv)

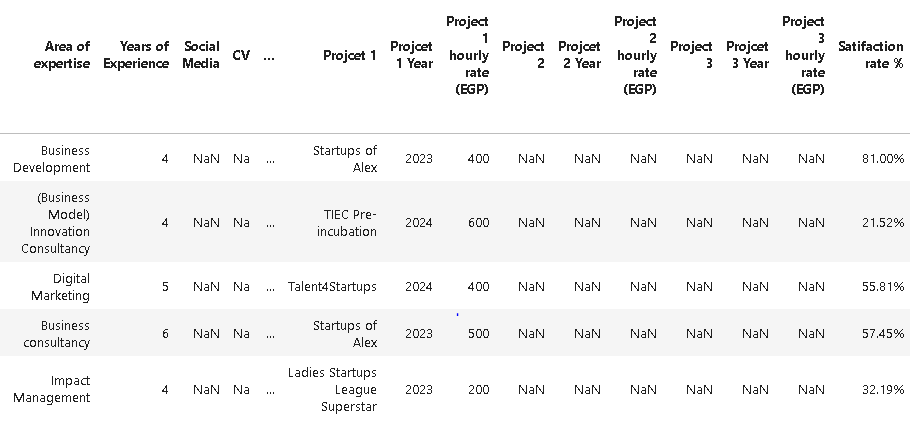
VALUES('Trainer31','fgdgsd','Female','Business Developer','Na','Business Development',4,'https://www.google.com/drive/')

**To delete a value**

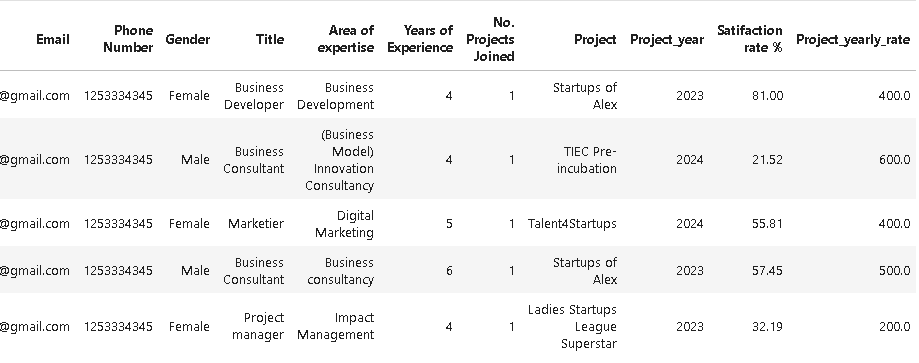
DELETE FROM Trainers

WHERE t\_code = 'Trainer01';

# **Exploratory Data Analysis(EDA):**



The data provided is not clean it contains columns with only null values like CV, and social media columns, and the format of the data is wide.



Dropping the CV, and social media columns that are not important in analysis, and handling missing values by imputing using mean, and median for numeric columns and mode for categorical values the data became clean,

Using the melt() method to unpivot the columns we convert the data from wide format to long format to be ready for analysis.

# **Pivot Tables:**

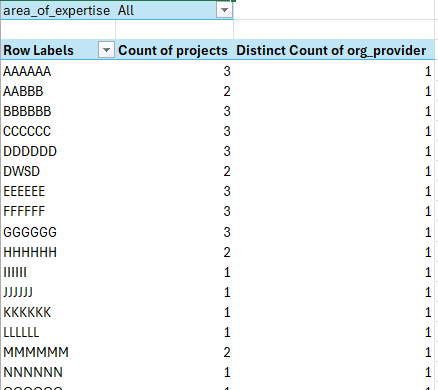
This section is divided into three parts.

* First, an analysis of all years of participation with ICEALEX is conducted.
* Second, an analysis of the data of the current year.
* Lastly, a report that shows the full data of the trainer.

## 6.1 Analysis and Reports overall:

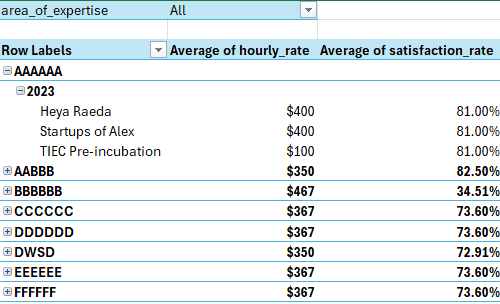
Here we provide an analysis of the whole data through the years.

### 6.1.1. Count of overall projects for each trainer



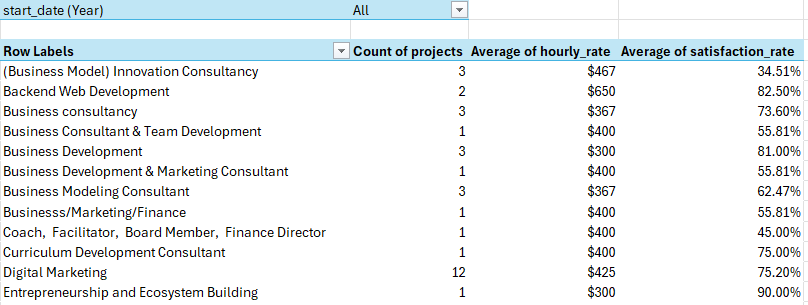
* This pivot table gives a summary of the **count of projects and the number of organizations**, each trainer worked in, throughout all years.
* You can use the filter to pick a specific category.

### 6.1.2. Records of hourly rate and satisfaction rate per trainers and projects throughout the whole years



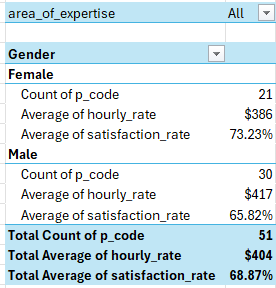
* This pivot table gives a summary record per each trainer through all years. It includes **the projects for each year, the average hourly rate, and the average satisfaction.**
* You can use the filter to pick a specific category.

### 6.1.3. Reports per subject



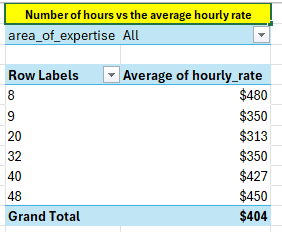
* This pivot table gives a summary record per category through all years. It includes **the projects for each year, the average hourly rate, and the average satisfaction.**
* You can use the filter to pick a specific year.

### 6.1.4. Data per Gender



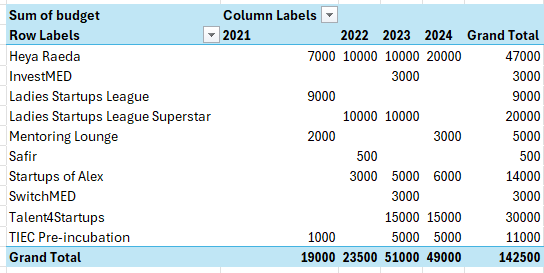
* This pivot table gives a summary record per gender through all years. It includes **the projects for each year, the average hourly rate, and the average satisfaction.**
* You can use the filter to pick a specific category.

### 6.1.5. Number of hours vs the average hourly rate



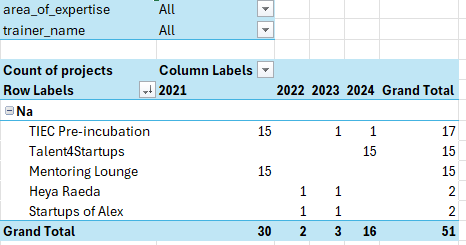
* This pivot table is to give a view of **the common number of hours vs the average hourly rate**.
* This is to give a quick expectation on the average hourly rate when making a contract with a trainer.
* You can use the filter to pick a specific category.(Can be used for new contracts).

### 6.1.6. Project data over years



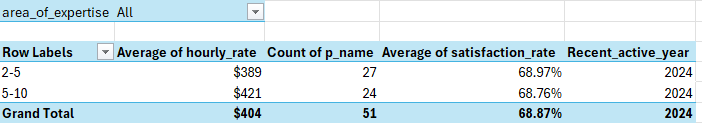
* This pivot table shows **the budget of the projects across the years.**

### 6.1.7. Organizations and projects



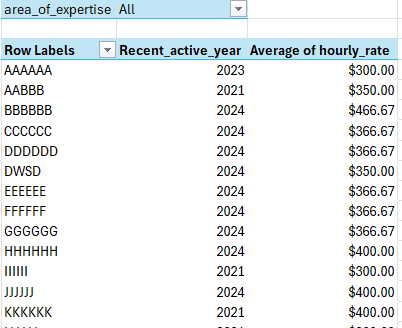
* This pivot table shows **the records of organizations, projects, and their count.**
* This can be filtered by category and trainer name.

### 6.1.8. View over experience level



* This pivot table shows **number of count, average hourly rate, and average satisfaction rate according to the level of experince through all active years**.
* Can be filtered based on category.(Can be used for new contracts)

### 6.1.9. Last active year and the hourly rate corresponding to this year for each trainer

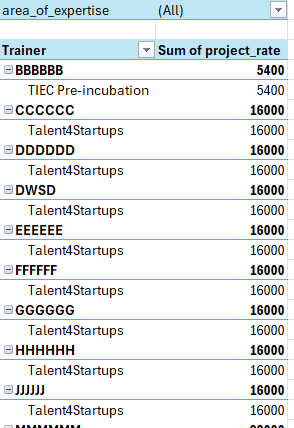


* This pivot table **shows the last year, a trainer was active with ICEALEX and the corresponding hourly rate in this year.**
* Can be filtered using category.

## 6.2 Analysis and Reports for the current year:

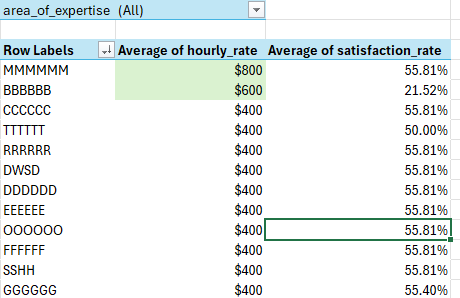
Here we provide an analysis of the data generated for the current year.

### 6.2.1.Trainer and project\_rate per project for the current year



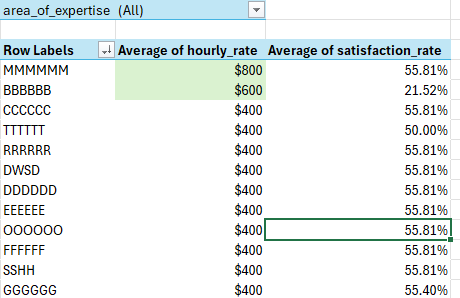
* This pivot table shows the **active trainers, current projects, project rate for each project and the total project rate for the *current year***.
* Can be filtered based on the category.

### 6.2.2.Average hourly rate for the current year per trainer



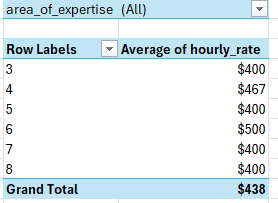
* This pivot table shows the **active trainers, their average current hourly rate, and the average satisfaction rate for the *current year***.
* Can be filtered based on the category.

### 6.2.3.Average hourly rate for the current year per trainer



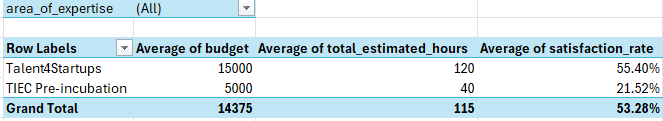
* This pivot table shows the **active trainers, their average current hourly rate, and the average satisfaction rate for the *current year***.
* Can be filtered based on the category.

### 6.2.4.Average hourly rate vs years of experience for the current year



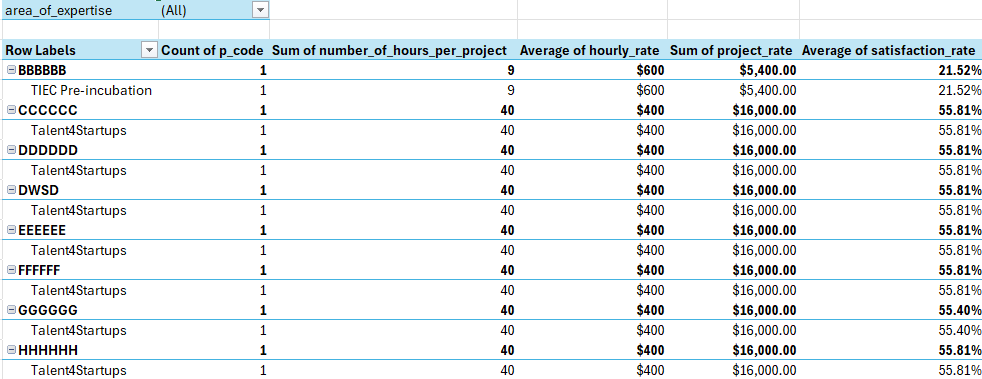
* This pivot table is to give a view over **the common number of hours vs the average hourly rate for the current year**.
* This to give a quick expectation on the average hourly rate when making a contract with a trainer.
* You can use the filter to pick a specific category.(Can be used for new contracts)

### 6.2.5.Project data for the current year



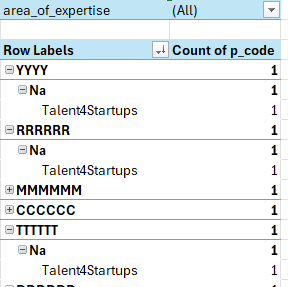
* This pivot table shows **the current active projects' data for the *current year*.**
* Can be filtered based on category.

### 6.2.6.Details of projects and trainers for the current year



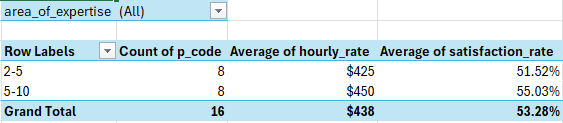
* This pivot table is to give **a summary for *current year's* projects fot each trainer**.
* You can use the filter to pick a specific category.

### 6.2.7.Trainer and organization



* This pivot table shows a **summary of the organizations and the projects the trainer participated in for the *current year*.**
* Can be filtered based on category.

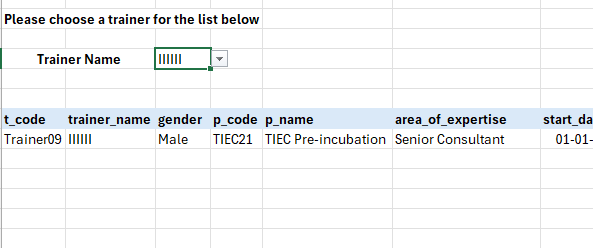
### 6.2.8.View over experience level for the current year



* This pivot table shows **number of count, average hourly rate, and average satisfaction rate according to the level of experience for *current year***.
* Can be filtered based on category.

## 6.3. A report of the trainer’s raw data

This report will be generated on picking a trainer from the list.



This view provides all the details for the trainer through the years.

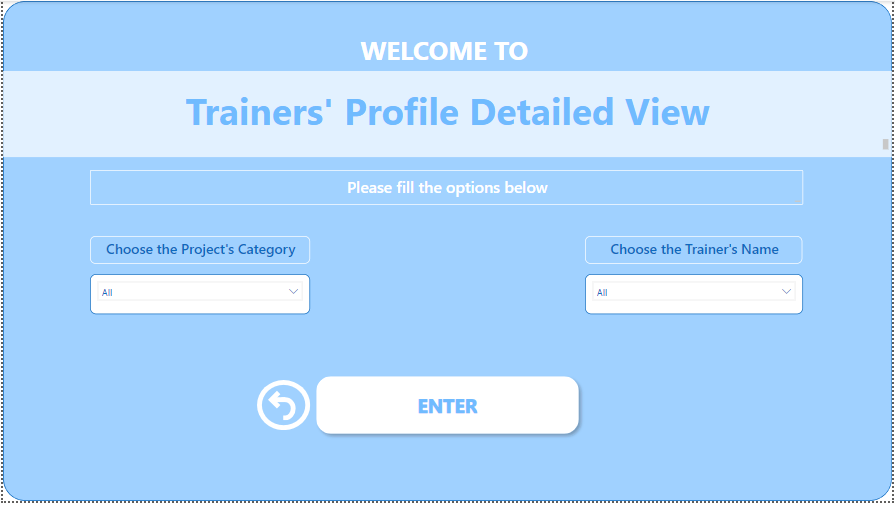
You need to pick the trainer's name from the dropdown list and all the trainer's details will appear.

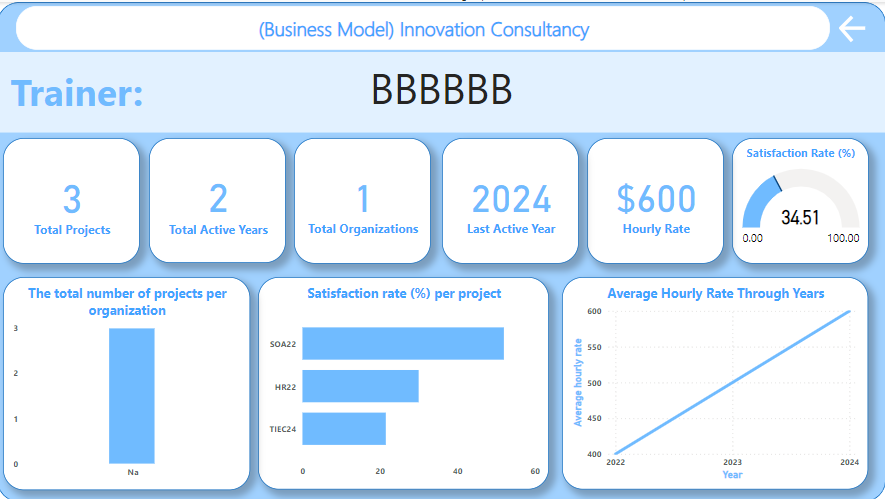
.

# **Analysis and Graphs:**

Here, we made a front page where you need to pick the area of expertise and the trainer’s name to access the detailed view of the chosen trainer.

There is a reset button to undo the filters.

After choosing from the slicers and pressing enter, a detailed dashboard including all the trainer’s details and the projects, the trainer participated in.



# **Recommendations:**

**These are simple recommendations, we thought of while working on the project, to enhance the efficiency of the whole process:**

* Fill in all fields when using the dashboard to avoid incorrect data.
* Categorize the area of expertise to subjects like(Data analysis, Marketing, etc..).
* For the project’s code entry, use the first letter of each word in the project name and the last two digits in the year to ensure its uniqueness.
* Satisfaction rate, we recommend adding an instant comment section if the rank applied doesn’t meet the target value. This will help get the cause of low rankings and help with improvement.
* Convert the years of expertise to experience levels instead.