PROJECT REPORT

Group No.: Group 07

Student Name: Dishita Jain

Executive Summary:

The principal objective of this project research and study is to design and implement a relational database which is industry ready for the application in the market of entrepreneurs and investors looking for business and who are constantly disappointed with the ideas or the funding they need

When we researched on this topic we find there is a huge amount of unstructured and noninformed data in the market which needs to be structures in order to make the most out of it and create some opportunities for both of the parties and it can only be done by implementing the relational database which helps in allocation of entities with the help of proper connection and also reduces data input process time by 50% and result in huge cost saving benefits across the industry. The database also implements a central analytics platform that has immense potential for analytics and fill this gap between the connections

The database of our company is designed in such a way that we have a list of investors, their shared interest about investment strategy and their desired capital. The database also consists of the name of those companies which are also interested in expanding and looking for some new ventures. Our database also has a distinguished and diversified projects which are at the final stage of matching with an investors interest. We have many options available for everyone which includes both and we believe in thorough confidentiality of our data

The database was modelled taking requirements of data fields required by us about investors, clients, their respective profiles and their requirements, along with input from prominent properties and all the down the line processes for input on the database. The EER and UML diagrams were modelled, followed by the mapping of the conceptual model to a relational model with the required primary and foreign keys. This database was then implemented fully MySQL and a prototype with two tables and two relationships were implemented on Mongo-DB NoSQL database using Studio 3T to study the feasibility of this database in a NoSQL environment

We have also tried to use Power-BI for better visualization of the data that we have collected and created the database with, and it includes the comparisons between different measures and prospects which is helpful in getting necessary insights and giving proper business-related guidance and decision-making points.

I. Introduction

As we know today many Undergraduates, Postgraduates and working individual are having a view of becoming an entrepreneur, in the belief that after completing their respective degree rather than working for someone else and earning some lucrative yet a fixed income, why don't we opt for a non-ending source of income and also to contribute in the betterment of society by creating jobs rather than finding one. The biggest problem any budding entrepreneur faces is the funding, whereas there are many investors as well as major companies who are looking for the innovation as well as to invest in young minds and talents. Here our company becomes the bridge between both the investors/ Companies and the Entrepreneurs, we believe in the saying Be Your Own Boss.:

When an entrepreneur comes with an idea, we have three categories of signup for the process. The first category denotes that it is a patent idea, second category is one with new idea yet didn't get any patent for that and the third category belongs to an innovation/enhancement in any existing idea. C.1 provides the most streamline process as the idea is a patent idea which is a win-win situation for both the investor as well as the entrepreneur, C.2 is an intricate process where we first rectify and check whether it is a new idea or is it already patent under any name or any similar kind of patent is available or not, whereas C.3 is the most sophisticated process where we need to introduce and include our team of experts for that idea and look for the positive chance of enhancement of the existing process/ idea is possible or not
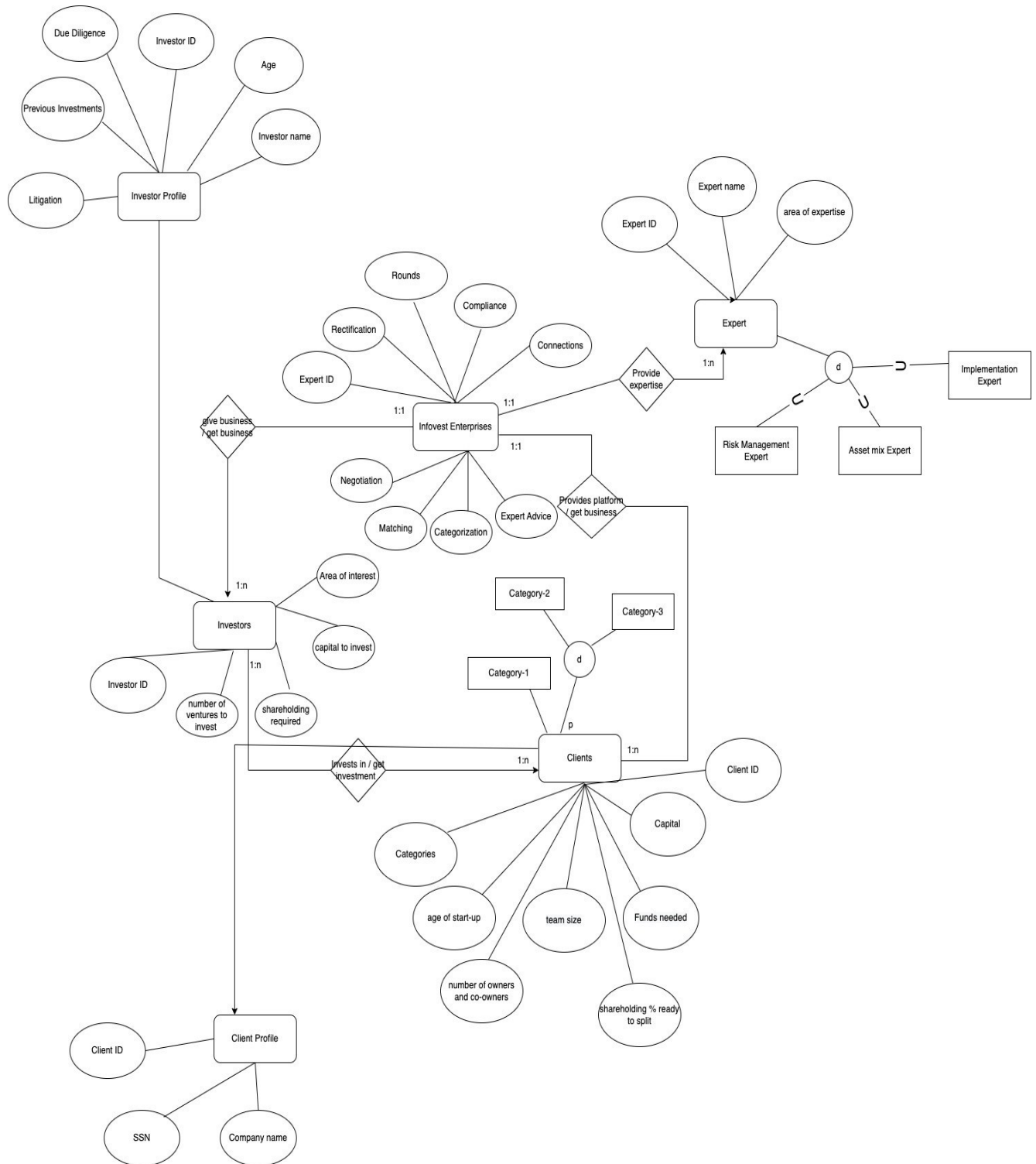
When any Investor or any Entrepreneur Visit the Facility or the online portal, we need some mandatory fields which should be filled by them before getting an appointment. For the investor the fields are area of interest, availability of funds, number of ventures to invest in, category type to invest in, % of shareholding required and respective others and for entrepreneurs the fields to be recorded are age of the startup, number or owners and coowners, Team size, Funds Needed, Category to apply in, % shareholding ready to split, any previous funding and respective others. Once we get the necessary data, we take 15 days to review and assessment of each case and then curate accordingly and execute the process according to the need of both, the process includes filtration, rectification, understanding needs, compliance, connections, expert advice, categorization, matching, InPerson meeting and negotiation.

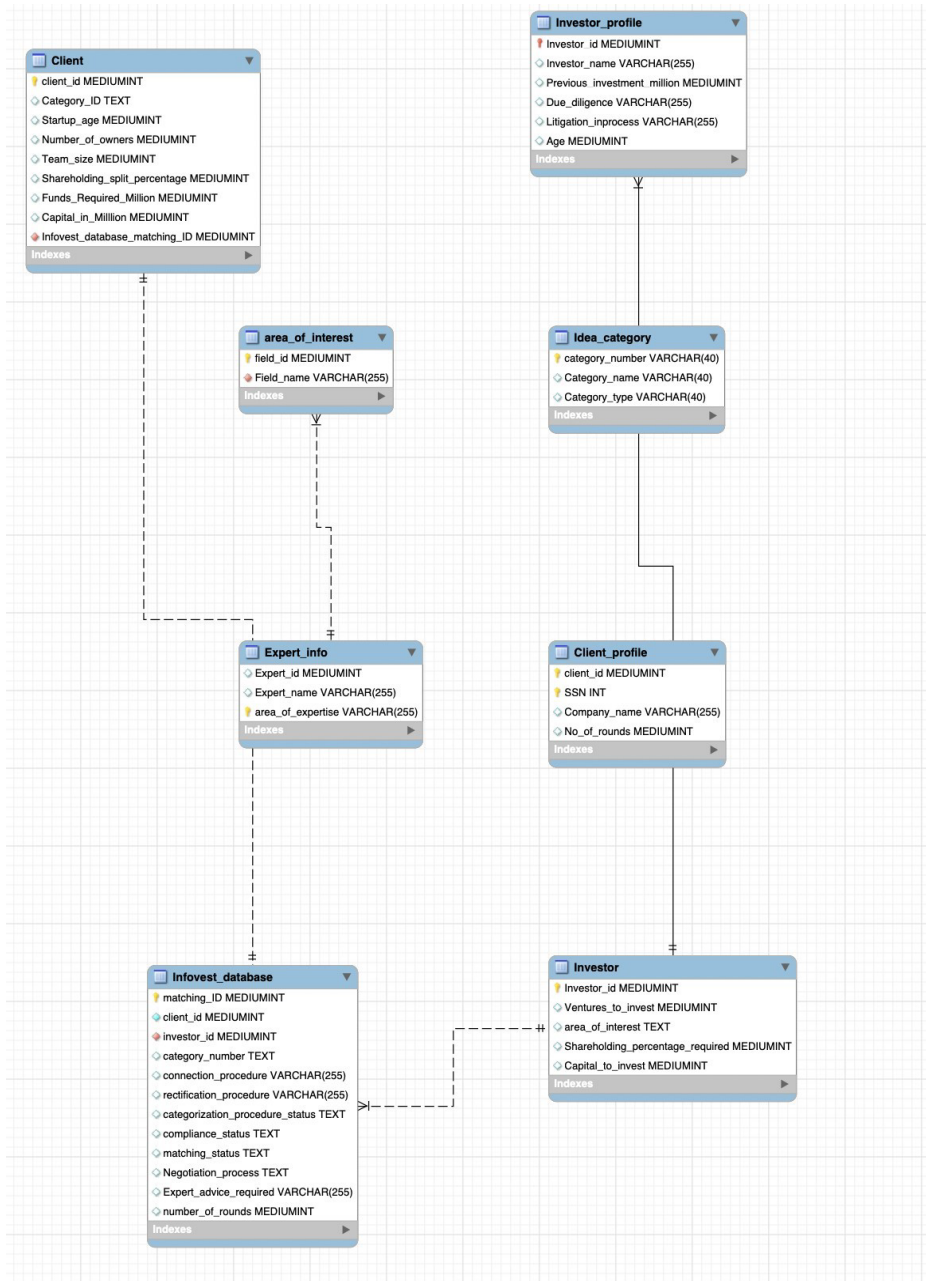**Principal requirements for enterprise model:**
- From the start of the process to the phase where both entities meet for the final signing up of the agreement, the names of the organizations shall not be disclosed
- An Investor can Invest in multiple ventures, but the entrepreneurs cannot seek/take any new investment for a cooling period of 6 months.

II. Conceptual Data Modeling

1. EER Model

2.UML Model

Client
- client_id MEDIUMINT
- Category_ID TEXT
- Startup_age MEDIUMINT
- Number_of_owners MEDIUMINT
- Team_size MEDIUMINT
- Shareholding_split_percentage MEDIUMINT
- Funds_Required_Million MEDIUMINT
- Capital_in_Million MEDIUMINT
- Infovest_database_matching_ID MEDIUMINT
- Indexes

Investor_profile
- Investor_id MEDIUMINT
- Investor_name VARCHAR(255)
- Previous_investment_million MEDIUMINT
- Due_diligence VARCHAR(255)
- Litigation_inprocess VARCHAR(255)
- Age MEDIUMINT
- Indexes

area_of_interest
- field_id MEDIUMINT
- Field_name VARCHAR(255)
- Indexes

Idea_category
- category_number VARCHAR(40)
- Category_name VARCHAR(40)
- Category_type VARCHAR(40)
- Indexes

Expert_info
- Expert_id MEDIUMINT
- Expert_name VARCHAR(255)
- area_of_expertise VARCHAR(255)
- Indexes

Client_profile
- client_id MEDIUMINT
- SSN INT
- Company_name VARCHAR(255)
- No_of_rounds MEDIUMINT
- Indexes

Infovest_database
- matching_ID MEDIUMINT
- client_id MEDIUMINT
- investor_id MEDIUMINT
- category_number TEXT
- connection_procedure VARCHAR(255)
- rectification_procedure VARCHAR(255)
- categorization_procedure_status TEXT
- compliance_status TEXT
- matching_status TEXT
- Negotiation_process TEXT
- Expert_advice_required VARCHAR(255)
- number_of_rounds MEDIUMINT
- Indexes

Investor
- Investor_id MEDIUMINT
- Ventures_to_invest MEDIUMINT
- area_of_interest TEXT
- Shareholding_percentage_required MEDIUMINT
- Capital_to_invest MEDIUMINT
- Indexes

III. Mapping Conceptual Model to Relational Model

Primary Key- <u>Underlined</u>   **Foreign Key- *Italicized***

- Investor- (**<u>Investor_id</u>**, Ventures_to_invest, area_of_interest, Shareholding_percentage_required,Capital_to_invest)

- Investor-profile
  (*Investor_id*,**<u>Investor_name,</u>** Previous_investment_million, Due_diligence,Litigation_inprocess,age)

- client (**<u>client_id</u>**, Category_id, Startup_age, Number_of_owners, Team_size, Shareholding_split_percentage, funds_Required_millions,Capital_in_million)

- client_profile - *Client_id*,**<u>SSN</u>**,Company_name,No_of_rounds

- Infovest database- (**<u>matching_ID</u>**, *client_id, investor_id, category_number*, connection_procedure, rectification_procedure, categorization_procedure_status, compliance_status,matching_status, Negotiation_process,Expert_advice_required,number_of_rounds)

- Expert_info- **<u>Expert_id</u>**,Expert_name,area_of_expertise,Expert_type

- Idea_category-**<u>category_number</u>**,category_name,category_type

- area_of_interest - **<u>field_id</u>**,field_name

IV. Implementation of Relation Model via MySQL and NoSQL

MySQL Implementations: -

Implementation of the relational model in MySQL

Q1- Query to get Investors who want shareholding percentage less than 8

SELECT *
  FROM Investor
  WHERE Investor_id IN (SELECT Investor_id
       FROM Investor
       WHERE
Shareholding_percentage_required < 8);

Investor_id	Ventures_to_inv...	area_of_interest	Shareholding_percentage_required	Capital_to_invest	
3	4	information technology	5	53	
7	4	Automobiles	7	74	
9	4	aeroneutics	6	91	
10	3	aeroneutics	2	86	
11	5	Insurance	2	82	
12	8	Insurance	6	88	
13	5	Realstate	5	79	
14	2	Realstate	3	69	
16	3	Financial Technologies	4	71	
17	6	Education	4	65	
20	7	Renewable Energy	6	24	
23	4	information technology	5	76	
24	7	information technology	5	4	
25	4	Pharmaceutical	6	15	
27	6	Automobiles	4	16	
29	4	aeroneutics	4	33	

Q2- Query to extract the Investors who have Previous investments more than 50 million and are not involved in any litigation.

SELECT
Investor_name,Previous_investment_million
, Litigation_inprocess
FROM Investor_profile
WHERE Previous_investment_million > 50
AND Litigation_inprocess = 'No';

Investor_name	Previous_investment_million	Litigation_inproc...
Macaulay Malone	74	No
Carson Mcbride	75	No
Ashton Rush	86	No
Russell Hutchinson	90	No
Kerry Singleton	73	No
Jared Briggs	80	No
Christopher Norman	89	No
Ciaran Blackburn	56	No
Cassady Andrews	89	No
Flavia Eaton	97	No
Carolyn Juarez	76	No
Uriel Hyde	72	No
Thomas Townsend	84	No
Arsenio Delacruz	99	No
Griffin Nash	64	No
Reagan Norton	63	No

Q3- Query to get Clients who require funds in million less than 100

```
SELECT *
  FROM Client
  WHERE client_id IN (SELECT
client_id
     FROM Client  WHERE Funds_Required_Million
```

client_id	Category_ID	Startup_age	Number_of_owne...	Team_size	Shareholding_split_percent...	Funds_Required_Million	Capital_in_Million
2	C1	1	2	130	56	11	1
15	C2	2	1	126	30	95	7
18	C3	5	5	35	55	99	4
20	C1	1	3	118	53	80	3
21	C2	8	2	118	56	39	2
22	C2	6	4	91	49	36	7
24	C3	7	4	82	29	6	7
25	C1	5	2	84	57	95	6
28	C2	3	3	25	37	21	1
36	C3	9	3	87	29	87	2
40	C2	7	3	130	60	91	8
42	C3	6	1	19	56	32	7
46	C2	5	2	119	48	23	0
47	C3	3	4	65	32	46	7
51	C2	3	2	129	37	71	4
52	C2	1	4	64	33	30	5

Q4- Query to get the total number of clients whose idea falls under category C2

```
SELECT count(*)
FROM Infovest_database
WHERE category_number ='C2';
```

count(*)
► 50

Result 1

Q5- Query to extract investor identity, required shareholding and due-diligence status from Investor and Investor Profile using Inner Join

```
SELECT
I.Investor_id,Investor_name,
Shareholding_percentage_required,D
ue_diligence FROM
Investor I
INNER JOIN Investor_profile Ip
ON I.Investor_id = Ip.Investor_id
ORDER BY
Investor_name,Shareholding_percent
age_required, Due_diligence;
```

Investor_id	Investor_name	Shareholding_percentage_required	Due_diligence
27	Adrienne Bradshaw	4	Yes
16	Ahmed Gay	4	No
40	Aiko Dominguez	2	Yes
127	Alana Hart	3	No
101	Alexander Lloyd	3	Yes
47	Ali O'donnell	10	Yes
52	Allistair Dodson	13	No
109	Anika West	14	No
85	Arden Juarez	7	No
129	Ariana Randall	4	Yes
108	Armando Holman	6	Yes
33	Arsenio Delacruz	15	Yes
4	Ashton Rush	9	No
87	Beatrice Madden	15	No
83	Benjamin Dean	5	Yes
30	Britanney Sweeney	14	No
7	Cade Edwards	7	Yes

Q6-Query to extract investor_id, capital to invest from investor where investor id matches with investor profile table

```
SELECT  Investor_id,
Capital_to_invest FROM Investor
WHERE EXISTS
(SELECT * FROM
Investor_profile
WHERE Investor.Investor_id =
Investor_profile.Investor_id);
```

Investor_id	Capital_to_invest
1	2
2	6
3	53
4	73
5	23
6	85
7	74
8	36
9	91
10	86
11	82
12	88
13	79
14	69
15	11
16	71
17	65

Investor 14

NoSQL Implementations: -

Implementation of the relational model in NoSQL

Q1-Query to find Investor_id, area_of_interest, Ventures_to_invest from Table investor

```
db.getCollection("Investor").find({},
{"Investor_id":"$Investor_id","area_of_interest":"$ar
ea_of_interest",
   "Ventures_to_invest" :
"$Ventures_to_invest","_id" : NumberInt(0)});
```

Investor > Investor_id		
Investor_id	area_of_interest	Ventures_to_invest
1	ROBOTICS	1
2	ROBOTICS	4
3	information tech	4
4	information tech	6
5	Pharmaceutical	2
6	Pharmaceutical	7
7	Automobiles	4
8	Automobiles	3
9	aeroneutics	4
10	aeroneutics	3
11	Insurance	5

1 document selected

Q2-Query to find expert information from Expert table.

```
db.getCollection("Expert_info").find({},
{"Expert_name":
1.0}).sort({"area_of_expertise": 1.0});
```

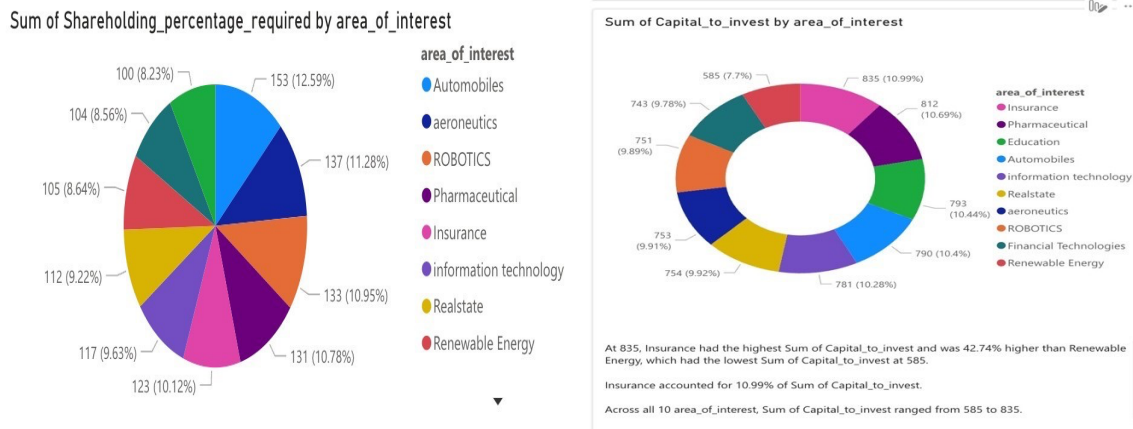| Expert_info > Expert_name	
_id	Expert_name
638bc992135b350ba4db66fd	Connor Morse
638bc992135b350ba4db66fe	Britanni Bender
638bc992135b350ba4db66ff	Cora Zimmerman
638bc992135b350ba4db6701	Elvis Mueller
638bc992135b350ba4db6702	Maile Le
638bc992135b350ba4db6705	Reagan Gonzalez
638bc992135b350ba4db6703	Phillip Stafford
638bc992135b350ba4db6704	Ulric Kent
638bc992135b350ba4db66fc	Shad O'donnell
638bc992135b350ba4db6700	Althea Lloyd

1 document selected

Q3-Query to find the count of number of clients from client table.

```
db.getCollection("client").aggregate
([{"$group" : {"_id" : { },"COUNT(*)" : {"$sum"
: NumberInt(1)}}},
 {"$project" : {"COUNT(*)" :
"$COUNT(*)","_id" : NumberInt(0)}}]);
```

client > COUNT(*)	
_id	COUNT(*)
	140

1 document selected

9

V. Database linked to Power-BI

Sum of Shareholding_percentage_required by area_of_interest

Sum of Capital_to_invest by area_of_interest

At 835, Insurance had the highest Sum of Capital_to_invest and was 42.74% higher than Renewable Energy, which had the lowest Sum of Capital_to_invest at 585.

Insurance accounted for 10.99% of Sum of Capital_to_invest.

Across all 10 area_of_interest, Sum of Capital_to_invest ranged from 585 to 835.

VI. Database Access via Python

The database is accessed using Python and visualization of analyzed data is shown below. The connection of MySQL to Python is done using MySQL. Connector, followed by cursor.excecute to run and fetchall from query, followed by converting the list into a data frame using pandas library and using matplotlib to plot the graphs for the analytics.
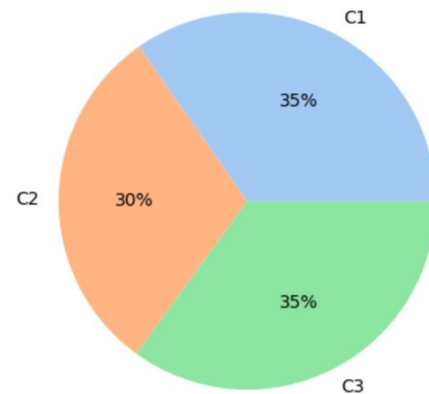
Graph 1: - Showing Capital to invest according to area of interest
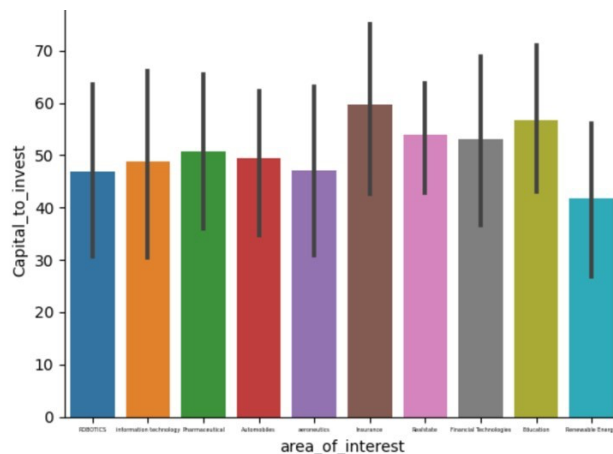
Graph 2: - Showing Funds required in million based on startup age
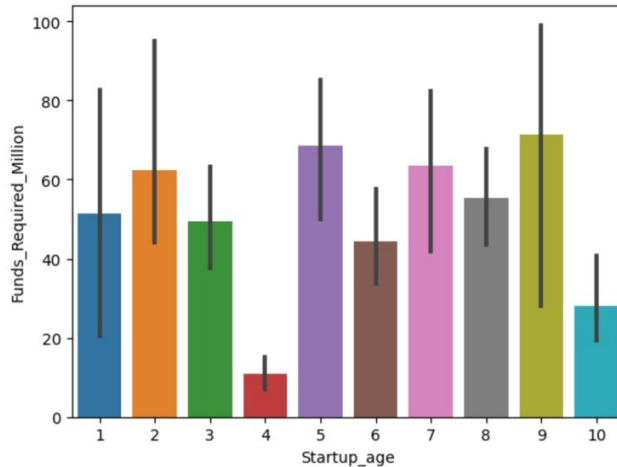
Graph 3: - Showing number of rounds each category idea must go through in order to finalize

VII. Summary and recommendation

The Infovest Enterprises database is an industry ready database and here we have a list of investors, their common interests in investing strategy, and the amount of capital they are seeking. The database also includes the names of businesses that are interested in growing and looking for new opportunities. The database also has the list of Experts, their types (which states what asset related expertise they have) and their area of expertise. Additionally, there are notable and diverse projects in our database that are close to meeting an investor's interest. We believe in complete confidentiality of our data and offer a variety of solutions for everyone.

When we conducted research on this subject, we discovered that there is a substantial amount of unstructured and uninformed data on the market that needs to be structured in order to make the most of it and open opportunities for both parties. This can only be accomplished by implementing a relational database, which aids in entity allocation with the help of proper connections and cuts the time required for data input by half.

Improvement on the database would be the implementation of Client filtering methods, the matching and negotiation process so that it can be quicker as every client and invertor

have their own terms and conditions and must be met for a successful matching. More study can be done as to how this unique relational database can be implemented. There are certain insights which would involve alternative investment fundraising and a reading room idea where investors can meet the client and talk about why they are interested in the company and gain each other's trust. For growth of the business, we would recommend the background, bankruptcy and security check to be considered a very important part of the process and events being held so that more and more people get to know about this.