T.E.R.A Corp Proposal

Theresa E. Randolph & Associates, LLC

IST 659 Final Project

September 2019

Data Analyst
Terrance Randolph
M.S Data Science (Student)
tmrandol@syr.edu
IST 659 Database Analytics

Table of Contents

Summary

Business Introduction

Internal Revenue Services known by their acronym IRS has been charged with collection, organizing and enforcing tax regulation for all citizens in the United States. Therefore, the burden of allocation and collection of revenue for millions of people and businesses is clearly too cumbersome for a single agency. Therefore, the collection and organization has been outsourced to tax preparer both large for example HR Block and Jackson Heweitt and small such as Randolph Tax, llc. These larger tax prepares more likely than not use proprietary software however, smaller companies use subscription based online services that allows them to compete technologically. Which brings to the forefront the only aspect hindering a smaller company from gaining scale and that is a strong database model and data analytics. Whereby, the SMB's (Small to Medium Businesses) can understand customer needs, revenue pitfalls and the ability to quickly adjust processes to accommodate efficiency.

Therefore, this project will focus on my mother's tax preparation side business that she wants to become a full-time venture after ten successful tax years. The reason being a recent increase in corporate or SMB's seeking out accountants to assist in filing multiple years of taxes and perhaps lowering business taxes owed. Furthermore, these better paying corporate clients require more tax preparation time than the individual clients but, the individual clients are still economically viable for the next few years. As a consequence, the peak tax season (Jan. - April) performance needs to increase in order to accommodate the almost year round corporate clients who taxes prep is longer and more laborious. Meaning, when corporate preparation spills into individual peak tax season the new smooth process can be automated and managed with money from both client categories secured and both clients pleased with Randolph Tax llc.

Business Problem

The business wants to consolidate their clients information into one standard format because their individual non - corporate clients are scattered between three tax preparation systems. Therefore, client consolidation would cut the cost of preparation software by 50% saving almost $200 a month. However, the question of how much efficiency would be gained with consolidation and will be the cost savings arose. Whereby, the structuring of a solid database seems to be one segment of a solution. However, the creation of a dashboard and data analytics will resolve the remaining segment of that problem. For instance, specific questions surrounding efficiency will drive the analysis of this project.

Efficiency Improvement Questions:

1) What is an adequate deadline for the client to turn in personal information/tax documents in full to tax preparer that will not impede profits? 1 day, 2 days or 7 days?
2) Due to arbitrary client rates, what are the best prices to charge when considering client refund amount, client tenure and time of preparation/filing?

Challenges

1) Creating a simple and easy to use process for entering the data into the system.
2) Migrating client from the aburitry text/phone call to a more structured system.

3) Searching for financially responsible methods of securing the clients data in the database or deciding not to store sensitive information.
4) Deciding whether or not the clients should have access to insights gained from their data. Even Though, a data use and consent form will be signed by the client prior to taxes being filed.

Expected Outcomes

1) Client: Organized database with automated and/or an easy method to input client information during preparation.
2) Client: Understanding what methods/processes will improve revenue and efficiency either current or new methods.
3) Analyst: Develop a functional database that will gain useful insights
    a) Geographic location, cost/ distance to tax preparer office
    b) Demographics/ target market, i.e. gender, age, income
    c) Peak business time within the tax season i.e. January, March

Conceptual Model

Analog Tables

Client Identification Table

Column Name	Type	Description
clientID	Integer 20	4 digit client unique identifier. FORMAT: [0001]
client_FirstName	Varchar 30	Client First Name FORMAT: [John]
client_LastName	Varchar 30	Client lastName FORMAT: [Doe]

Client Information Table

Column Name	Type	Description
clientID	Integer 20	4 digit client unique identifier. FORMAT: [0001]
Document_Retreival_Method	Varchar 20	How were the w-2 received FORMAT: [Email, Drop-off]
Refund_Amount	Integer 30	Refund return or deficit, negatives are deficits FORMAT: [-5000, 5000]
Type_of_Filing	Character 50	Client type of filing FORMAT: [1040/Sch C, 1120 etc...]
Filing_Status	Character 50	Client IRS filing status FORMAT: [Single, Head of HouseHold etc...]
Client_Income	Integer 20	Income range within 5k FORMAT: [0 - 10,000,000]
tax_year	Integer 10	The year taxes were filed FORMAT: [2018, 2019]
Billed_Amount	Integer 20	The full/total amount billed to the client FORMAT: [0 - 10,000,000]
Payment Plan	Integer 20	Boolean of whether or not client is on a payment plan FORMAT: [Yes, No]
Installment_Amount	Integer 20	The payments are assumed to be even, therefore, enter single payment amount FORMAT: [0 - 10,000]
Date_of_Filing	Datetime 10	The start date when tax preparer submits the W-2 FORMAT: [04/14/2018]

Client Demographic Table

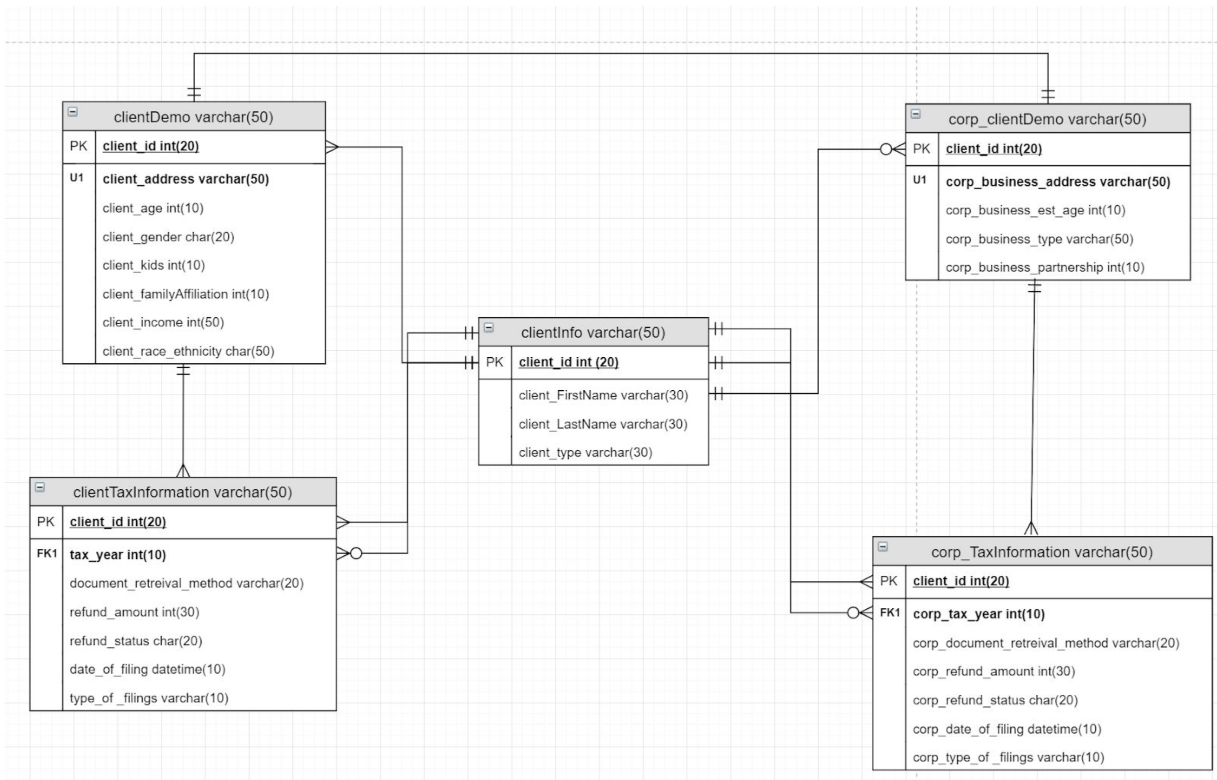
Column Name	Type	Description
clientID	Integer 20	4 digit client unique identifier. FORMAT: [0001]
Client_Age	Integer 10	AgeFORMAT: [01]
Client_Kids	Integer 10	Number of kids from 0-1 FORMAT: [0 - 10]
Client_Race	Character 30	Client race FORMAT: [AA 'Afriacan American, WA 'White American, AL 'African Liberian']
Client_Gender	Character 30	Male or Female: M,F FORMAT: [M 'Male', F 'Female', O 'Opted out']
Client_Address	Varchar 50	Client City and State FORMAT: [1250 Address etc...]

Client Corporate Specialty Table

Column Name	Type	Description
clientID	Integer 20	4 digit client unique identifier. FORMAT: [0001]
Corp Business Age	Integer 10	Business establishment date FORMAT: [1999,2019]
Corp Business Type	Varchar 50	Business type, llc,sole proprietorship etc.. FORMAT: [llc, corp, etc...]

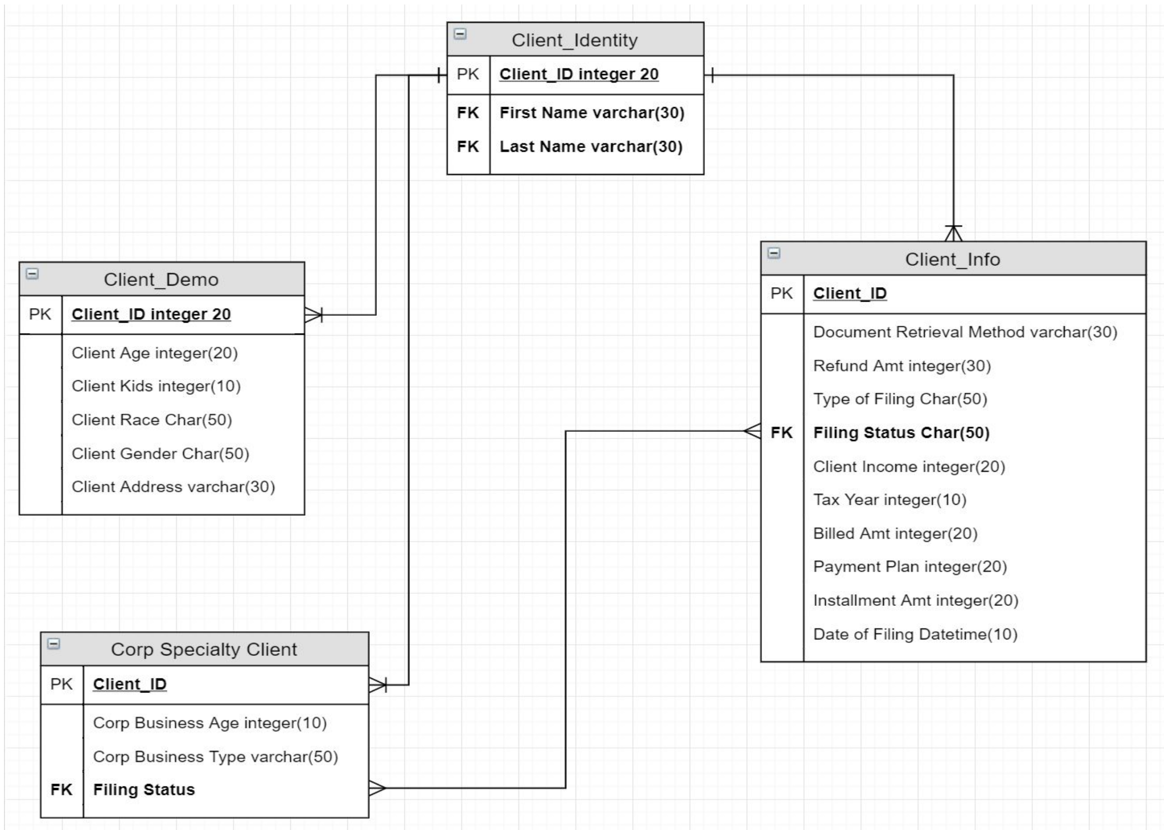
Entity Relational Diagrams

Initial ERD



Final ERD

(adjusted to fit business better)



Attributes Relationships

- The **ClientID** is unique to **First_Name** and **Last_Name**
- The **Filing Status** can be applied to Individual clients as well as Corporate clients
- There can be many **Tax_Years** associated with one client and corporate client
- There is a many to many **Client Demo** to **Client Info** due to a client demographics changing via increasing family size which impacts Filing Status.
- **ClientID** has a one to many relationships with **ClientDemo, ClientInfo and Corp Specialty Table**.

Analysis & Discovery

Final Project Deliverables

- SQL Database populated with data (imported excel)

- Google Sheet with data entry dialog form (Google cloud based)
  - Created a DBMS with google forms and backend coding
  - User can enter data that is instantly saved to spreadsheet
  - User entry will automatically update dashboard with new data

- Dashboard
  - Dashboard gives a snapshot of the business standing
  - Established user friendly filters; Filing Date & Client Name
  - Interactive insights of Billed amount, Client Refund etc…
  - Various graphs and plots answering useful questions
    - Refund to Income Ratio
    - Annual Billed Amount etc...

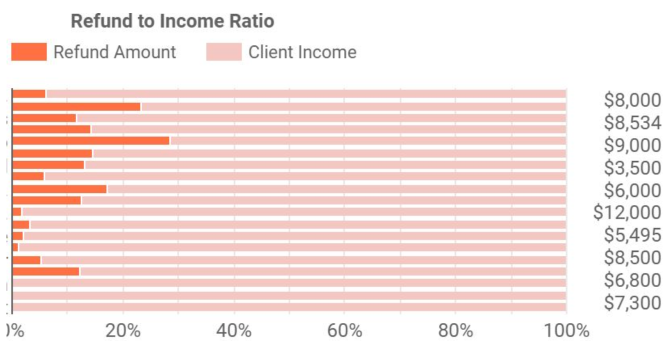
Question Answered

- What is TERA Corp client refund to income ratio and average?
  - **Refund Avg:** $2,913
  - **Billing Avg:** $155
  - **Income Avg:** $46,947
  - **Billing to Refund Avg :** 5%

```
 SELECT ROUND(AVG( Refund_Amt ),0) AS Refund_Avg,
        ROUND(AVG( Billed_Amt ),0) AS Billed_Avg,
        ROUND(AvG( client_Income ),0) AS income_Avg
FROM dbo.Corp
SELECT
     ROUND(AVG( Billed_Amt ) / AVG( Refund_Amt ) * 100,0) AS
Billed_pct_Refund_avg
FROM dbo.Corp
```

	Refund_Avg	Billed_Avg	income_Avg
1	2913	155	46947

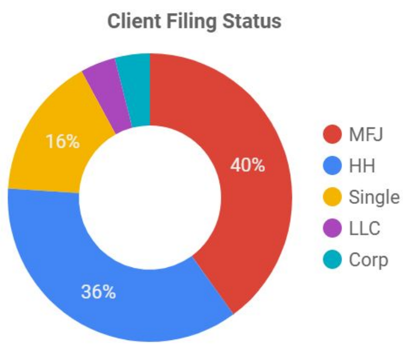
	Billed_pct_Refund_avg
1	5



- How many clients are in a specific filing type?
  - T.E.R.A Corp recently embarked on the corporate aspect of U.S Tax filing therefore, percentage of clients are low in that category. However, they will focus their non-corporate aspect of the business on Married couples and Single parents (HH).
  - **Married Filing Joint:** 40%
  - **Head of HouseHold:** 36%
  - **Single:** 16% (zero dependents)
  - **LLC & Corp :** 1% Respectively

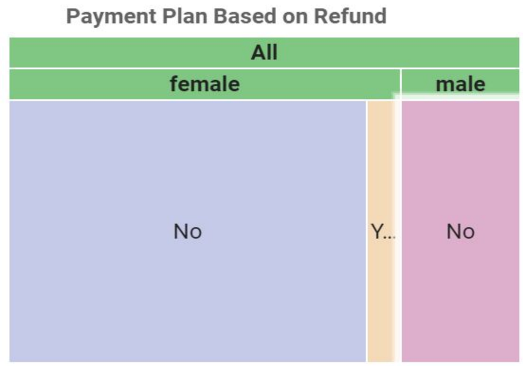
```
SELECT
      Filing_Status,
      COUNT(Filing_Status) AS CountOf_Status
FROM dbo.Corp
Group by Filing_Status
```

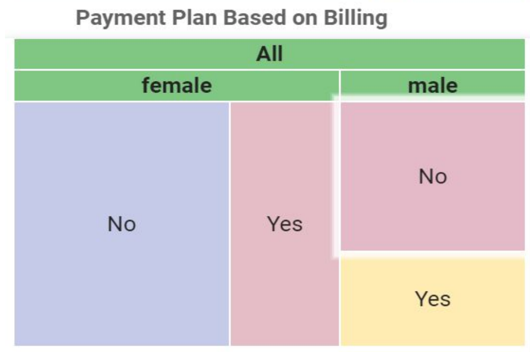
	Filing_Status	CountOf_Status
1	Corp	3
2	HH	25
3	LLC	3
4	MFJ	28
5	Single	8



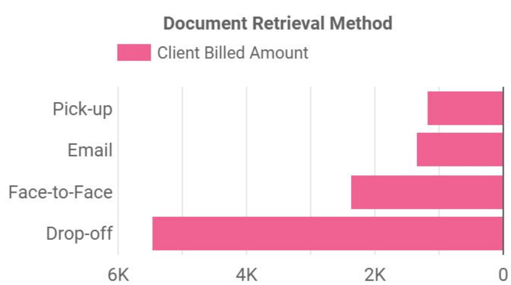
- Do clients with kids receive a higher refund? Is TERA Corp embracing that high children market?
    - Yes, the business is focusing their efforts on acquiring new clients with larger families due to that particular family initiative to complete taxes early. Also, they will receive higher refunds therefore, billing is higher and the work is not much more for that client type.

- Are payment plans increasing revenue or decreasing revenue?
    - Clients on a payment plans are about a 3rd of the female client population which would not necessarily cause a problem. However, the installments are four or more on average causing revenue to be unstable do to incompletion of payment plan. Therefore, the analyst (myself) suggested to T.E.R.A Corp that if payment plans are absolutely necessary keep installments below three. In order to ensure full and timely payment from clients.





- Which document retrieval methods are more successful?
  - Drop-off yields the highest amount Billed however, combined with other virtual methods of delivery Face to Face encompases a little over a quarter of billing. Therefore, migrating in-person clients to a digital method (phone call allowed) my lower overhead, which inturn will increase revenue.



Business Development Prospects

Creating a more robust UI (User Interface) for T.E.R.A Corp management and an **emailable form for the clients** to fill out and their responses will update the database and dashboard.Once, more data is collected T.E.R.A Corp can predict which clients IRS forms will be **flagged** and quickly estimate what **refund** would be. Therefore, the dashboard will give T.E.R.A Corp annual insights into clients individual and aggregated refund, income and filing dates. That of which can shift how efficient the business can adjust to clients. Consequently, a website that will read their **documents OCR** (Optical Character Recognition) and Text Mining features that will allow the business to keep the overhead low but the skills and security is challenging.

Dashboard & Google Sheets

Dashboard

Justification

- Intuitive navigable insights and filters for T.E.R.A Corp
- Digestible visualizations

Google Sheets

Justification

- T.E.R.A Corp requested cheap/free cloud services
- Clients Familiarity with google sheets

```
// When Sheet is opened the dialog pop-up is triggered
//     google forms iframe (embedded) will show.
function onOpen() {
  SpreadsheetApp.getUi() // Or DocumentApp or SlidesApp or FormApp.
      .createMenu('Dialog')
      .addItem('Open', 'openDialog')
      .addToUi();
}

function openDialog() {
  var html = HtmlService.createHtmlOutputFromFile('Tax HTML Final');
  SpreadsheetApp.getUi() // Or DocumentApp or SlidesApp or FormApp.
      .showModelessDialog(html, 'T.E.R.A Corp & Associates, LLC');
//.showModalDialog(html, 'T.E.R.A Corp & Associates, LLC');
}
```

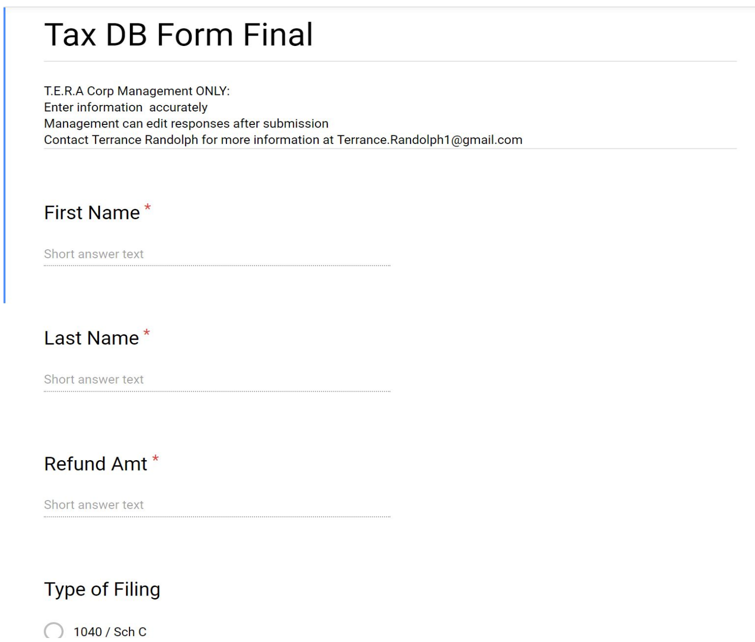
Google Forms

Justification

- Easily integratable GUI for google sheets

Table & Form Sample

Google Forms



Tax DB Form Final

T.E.R.A Corp Management ONLY:
Enter information accurately
Management can edit responses after submission
Contact Terrance Randolph for more information at Terrance.Randolph1@gmail.com

First Name *

Short answer text

Last Name *

Short answer text

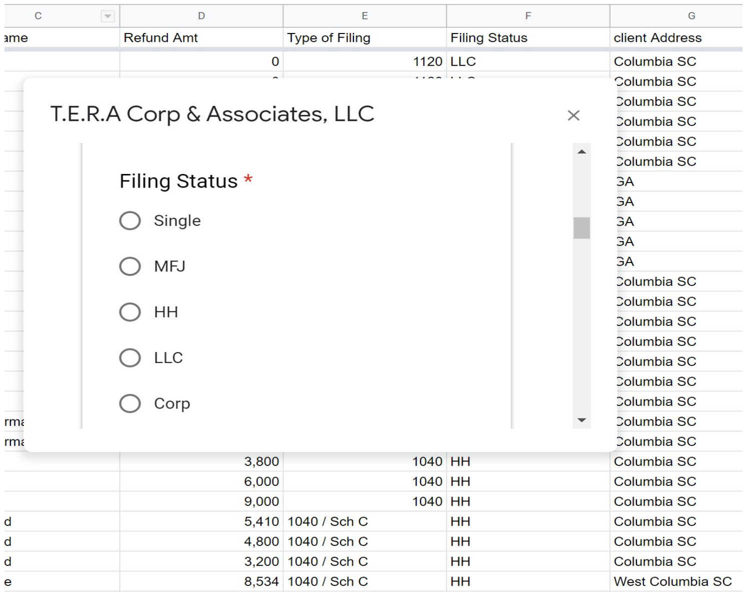
Refund Amt *

Short answer text

Type of Filing

1040 / Sch C

GoogleSheets Table



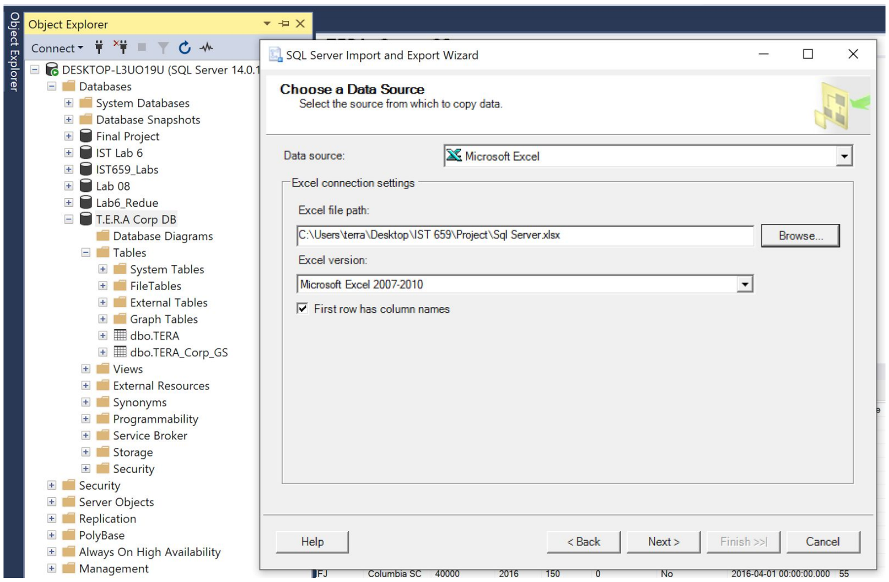
Visualizations

Dashboard Sample Link:

https://datastudio.google.com/s/vZYCpKVsTNE

SQL Table Creation

Table was originally created in SQL Server then transferred into google sheets for clients ease of use. Therefore, an import from excel that was downloaded from google sheets was necessary. Sample image below.



Original SQL Table Creation

```
/*
                    Terrance Randolph
                Final Project
                  Proposal 1

*/
---------------------------------------
--- Client Client information Table ---
---------------------------------------
CREATE TABLE clientInfo (
      --Columns For The Client information Table
```

```sql
	clientID int identity not null,
	client_FirstName varchar(30),
	client_LastName varchar(30),
	client_type varchar(30) not null,
	-- Constraints For The Client information Table
	CONSTRAINT PK_clientInfo PRIMARY KEY (clientID))

--- Created! Client information Table ---

---------------------------------------
--- Client Client Demographic Table ---
---------------------------------------
CREATE TABLE clientDemo (
	--Columns For The Client Demographic Table
	clientID int identity not null,
	client_address varchar(50),
	client_age int,
	client_kids int default 0,
	client_familyAffiliation int default 0,
	client_income int not null,
	client_race_ethnicity char(50) default 'UNK',
	-- Constraints For The Client Demographic Table
	CONSTRAINT PK_clientDemo PRIMARY KEY (clientID),
	CONSTRAINT U1_clientDemo UNIQUE(client_address))

--- Created! Client Demographic Table ---

-------------------------------------------
--- Client Client Tax Information Table ---
-------------------------------------------
CREATE TABLE clientTaxInformation (
	--Columns For The Tax Information Table
	clientID int identity not null,
	tax_year int not null,
	document_retrevial_method varchar(20) not null,
	refund_amount int not null,
	refund_status char(20) not null,
	date_of_filing datetime not null,
```

```
      type_of_filing varchar(10) not null default 'W2',
      -- Constraints For The Client Tax Information Table
      CONSTRAINT PK_clientTaxInformation PRIMARY KEY (clientID),
      CONSTRAINT FK_clientTaxInformation FOREIGN KEY (tax_year)
References clientInfo(clientID))

--- Created! Client Tax Information Table ---

----------------------------------------------
--- Client Corporate Tax Information Table ---
----------------------------------------------
CREATE TABLE corp_TaxInformation (
      --Columns For The Corporate Tax Information Table
      clientID int identity not null,
      corp_tax_year int not null,
      corp_document_retrevial_method varchar(20) not null,
      corp_refund_amount int not null,
      corp_refund_status char(20) not null,
      corp_date_of_filing datetime not null,
      corp_type_of_filing varchar(10) not null,
      -- Constraints For The Corporate Tax Information Table
      CONSTRAINT PK_corp_TaxInformation PRIMARY KEY (clientID),
      CONSTRAINT FK_corp_TaxInformation FOREIGN KEY (corp_tax_year)
References clientInfo(clientID))

--- Created! Corporate Tax Information Table ---

------------------------------------------
--- Client Corporate Demographic Table ---
------------------------------------------
CREATE TABLE corp_clientDemo (
      --Columns For The Corporate Client Demographic Table
      clientID int identity not null,
      corp_business_address varchar(50),
      corp_business_est_age int,
      corp_business_type varchar(50),
      corp_businss_partnership int not null default 0,
      -- Constraints For The Corporate Demographic Table
```

```
     CONSTRAINT PK_corp_clientDemo PRIMARY KEY (clientID),
     CONSTRAINT U1_corp_clientDemo UNIQUE(corp_business_address))

--- Created! Corporate Demographic Table ---
```