DATABASE NORMALIZATION PROJECT

Data Normalization is the process of removing redundancies and improving data integrity of a database. This will prevent different anomalies that may arise from poor database design. Ideally, a database should reach the Third Normal Form.
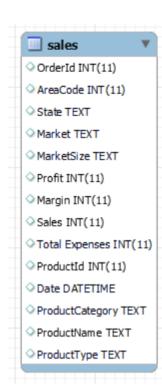
In this project, a store named **Hello Coffee** has their sales data stored in a MySQL database. However, the database only contains a single table, containing attributes that represent other entities aside from sales.

hello_coffee database: sales table

This database, contains only one single table named sales. This was the table prior to normalization.

- This table is a standalone table inside the database.
- The table adheres to the 1NF. However, there are non-key attributes that are partially dependent only to the composite keys OrderId, AreaCode, ProductId.
- Further design improvement is needed to reach 2NF and 3NF.

OrderId	AreaCode	State	Market	MarketSize	Profit	Margin	Sales	Total Expenses	ProductId	Date	ProductCategory	ProductName	ProductType
3412	206	Washington	West	Small Market	38	77	130	39	2	2010-01-01 00:00:00	Coffee	Columbian	Regular
3413	206	Washington	West	Small Market	39	81	150	42	6	2010-01-01 00:00:00	Espresso	Decaf Espresso	Decaf
3414	206	Washington	West	Small Market	115	186	310	71	8	2010-01-01 00:00:00	Herbal Tea	Chamomile	Decaf
3415	213	California	West	Major Market	203	312	546	109	6	2010-01-01 00:00:00	Espresso	Decaf Espresso	Decaf
3416	216	Ohio	Central	Major Market	67	116	200	49	12	2010-01-01 00:00:00	Tea	Earl Grey	Regular

sales
- OrderId INT(11)
- AreaCode INT(11)
- State TEXT
- Market TEXT
- MarketSize TEXT
- Profit INT(11)
- Margin INT(11)
- Sales INT(11)
- Total Expenses INT(11)
- ProductId INT(11)
- Date DATETIME
- ProductCategory TEXT
- ProductName TEXT
- ProductType TEXT

After Normalization

After normalizing hello_coffee, this is now the new EER diagrams. Each table in this database now only represents a single entity.