

Final Design Report



Rosetta

DROP TABLE Groups;

Database Modeling, Design and Implementation

CPE 366 Section 01

Brian Holland bholla02@calpoly.edu

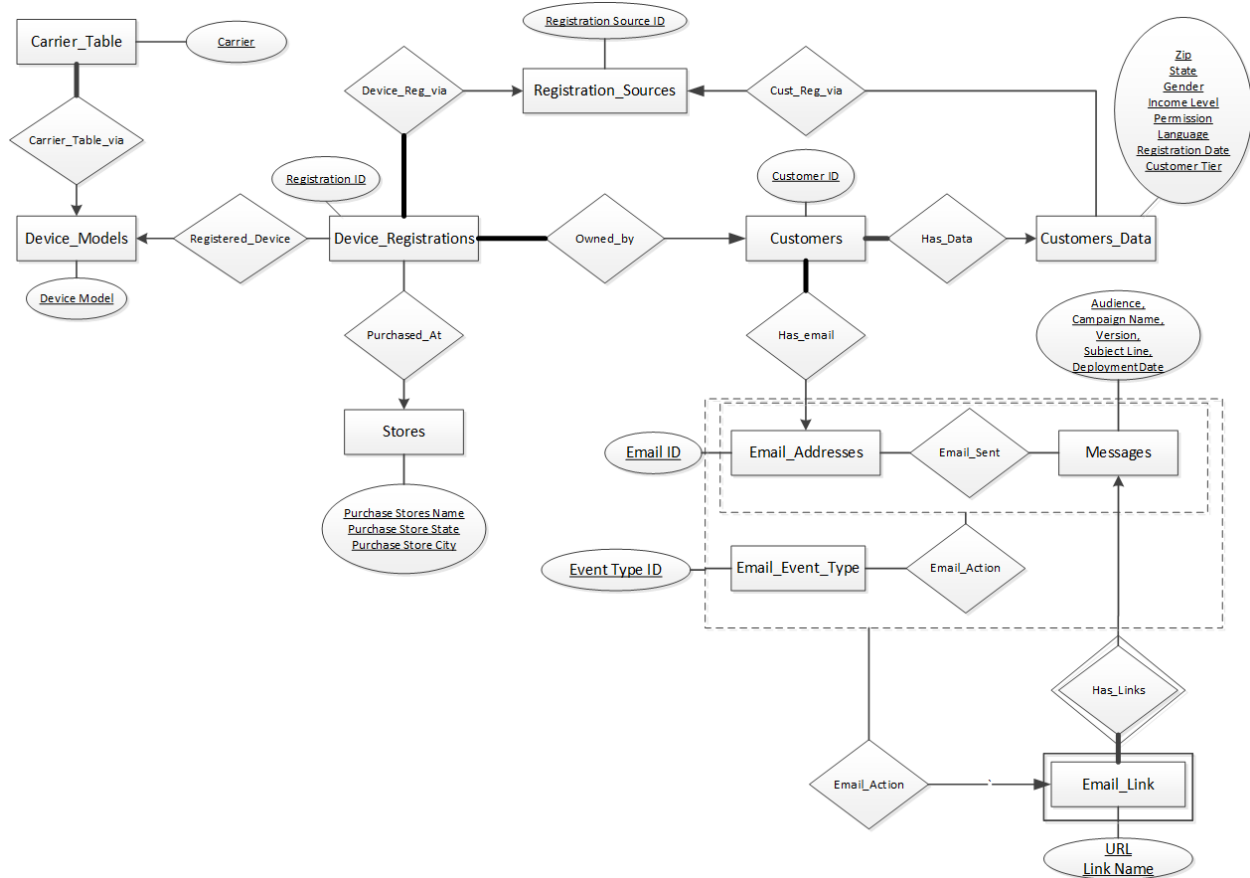
Michael Slevin mslevin@calpoly.edu

Stefan Bonilla sabonill@calpoly.edu

Scott Tucker srtucker@calpoly.edu

Jenna Murphy jmurph14@calpoly.edu

E-R Model for Operational Data Store



Relational Model for the Operational Data Store

Entities

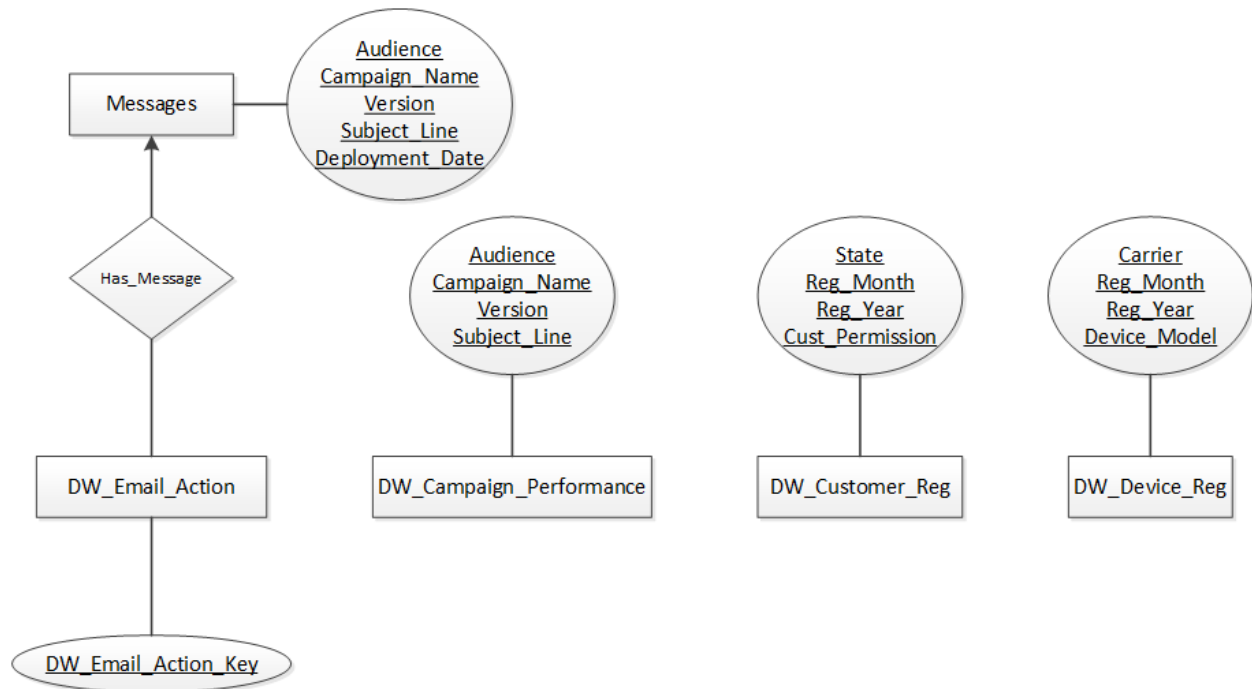
- Registration_Sources(Registration Source ID (String), Registration Source Name (String))
- Customers(Customer ID (String))
- Email_Addresses(Email ID (String), Email Domain (String))
- Device_Models(Device Model (String), Device Name (String), Device Type (String), Number of Registrations (Int))
- Stores(Purchase Store Name (String), Purchase Store State (String), Purchase Store City (String))
- Device_Registrations(Registration ID (String), Registration Date (Date), Ecomm Flag (Boolean), Purchase Date (Date), Serial Number (String))
- Email_Campaign(Audience (String), Campaign Name (String), Version (String), Subject Line (String), Deployment Date (Date))
- Email_Event_Type(Event Type ID (String), Event Type Name (String))
- Email_Link(URL (String), Link Name (String))

- Carrier_Table(Carrier(String))
- Customers_Data(Zip (Int), State (String), Gender (String), Income Level (String), Permission (Boolean), Language (String), Registration Date (Date), Customer Tier (String))

Relationships

- Cust_Reg_via(Customers, Registration_Sources)
 - One-to-many
 - Required
- Has_email(Customers, Email_Addresses)
 - One-to-many
 - Required
- Owned_by(Device_Registrations, Customers)
 - One-to-Many
 - Required
- Registered_Device(Device_Registrations, Device_Models)
 - One-to-many
 - Optional
- Carrier_Table_via(Carrier_Table, Device_Models)
 - One-to-many
 - Required
- Device_Reg_via(Device_Registrations, Registration_Sources)
 - One-to-many
 - Required
- Purchased_At(Device_Registrations, Stores)
 - One-to-many
 - Optional
- Email_Sent(Email_Addresses, Email_Campaign)
 - Many-to-many
 - Attributes:
 - Deployment ID (String)
- Has_Data(Customers, Customers_Data)
 - One-to-Many
 - required
- Email_Action(aggregate Email_Sent(Email_Addresses, Email_Campaign), Email_Event_Type)
 - Many-to-many
 - Attributes:
 - Event Date (Date)

E-R Model for the Data Warehouse



Relational Model for the Data Warehouse

Entities

- DW_Email_Action(DW_Email_Action_Key (Int), Message_Key (Int), Event_TypeName (String), Count (Int))
- DW_Campaign_Performance(Audience (String), Campaign_Name (String), Version (String), Subject_Line (String), numClicks (Int), numComplaint (Int), numOpened (Int), numDelivered (Int), numUnsubscribe (Int), numBounce (Int), numNotSent (Int))
- DW_Customer_Reg(State (String), Reg_Month (Int), Reg_Year (Int), Cust_Permission (Int), Count (Int))
- DW_Device_Reg(Carrier (String), Reg_Month (Int), Reg_Year (Int), Device_Model (String), Count (Int))

Relationships

- Has_Message(Messages, DW_Email_Action)
 - One-to-Many

ETL Process Description

We first created four temporary tables corresponding to the Account Registration Data, Device Registration Data, Email Event Data, and Device Data tables given to us in the project specifications. We loaded all these temporary tables with the customer data via a load SQL script using LOAD DATA INFILE statements. After loading the given data, we created our own

tables representing our operational data store for the project. By using another SQL script, we then transformed our temporary tables holding the loaded data into our operational data store by using INSERT statements. These statements inserted data into our operational data store by querying relational tables from joining our temporary tables with previously transformed tables through SELECT statements.

Data Dictionary

Entity	Attribute	Data Type	Definition
Registration_Sources	Reg_Source_Key	INT	Auto Increment key
	Registration Source ID	VARCHAR(MAX)	ID of registration source for customers and devices. Taken from flat data.
	Registration Source Name	VARCHAR(MAX)	Name of registration source for customers and devices. Taken from flat data.
Customers	Customer_Key	INT	Auto Increment key
	Customer ID	VARCHAR(MAX)	Unique ID to identify customers. Taken from flat data.
	Reg_Source_Key	INT	Foreign key to Registration_Sources that identifies where the customer was registered from.
Customer_Data	Customer_Data_Key	INT	Auto Increment key
	Zip	CHAR(5)	ZIP code from customer registration. Customer-reported, taken from flat data.
	State	VARCHAR(MAX)	State from customer registration. Customer-reported, taken from flat data.
	Gender	VARCHAR(MAX)	Gender from customer registration. Customer-reported, taken from flat data.
	Income Level	VARCHAR(MAX)	Income level from customer registration. Customer-reported, taken from flat data.
	Permission	VARCHAR(MAX)	Permission level of registered customer. Set by Brand X at time of registration. Taken from flat data.
	Language	VARCHAR(MAX)	Language from customer registration. Customer-reported, taken from flat data.
	Registration Date	DATETIME	Date of registration. Set by Brand X at time of registration. Taken from flat data.
	Customer Tier	VARCHAR(MAX)	Tier of customer ("Mid", "VIP"). Set by Brand X at time of registration. Taken from flat data.
	Customer_Key	INT	Foreign key to Customers table.
Email_Addresses	Email_Key	INT	Auto Increment key

	Email ID	VARCHAR(MAX)	Email of customer from Account Registration Data. Customer-reported, taken from flat data.
	Customer_Key	INT	Foreign Key to Customers table.
	Email Domain	VARCHAR(MAX)	Email Domain from Account Registration Data. Domain of the customer email. Customer-reported, taken from flat data.
Carriers	Carrier_Key	INT	Auto Increment key
	Carrier_Name	VARCHAR(MAX)	Carrier of the device from Device Data. Taken from flat data.
Device_Models	Device_Key	INT	Auto Increment key
	Device Model	VARCHAR(MAX)	Model of device from Brand X. Taken from flat data.
	Device Name	VARCHAR(MAX)	Name of device from Brand X. Taken from flat data.
	Device Type	VARCHAR(MAX)	Type of device from Brand X (tablet, phone, etc). Taken from flat data.
	Carrier_Key	INT	Foreign key to Carriers table to identify what carrier the device uses.
Stores	Store_Key	INT	Auto Increment key
	Purchase Store Name	VARCHAR(MAX)	Name of the store where the device was purchased. Reported by customer, taken from flat data.
	Purchase Store State	VARCHAR(MAX)	State of the store where the device was purchased. Reported by customer, taken from flat data.
	Purchase Store City	VARCHAR(MAX)	City of the store where the device was purchased. Reported by customer, taken from flat data.
Device_Registrations	Registration_Key	INT	Auto Increment key
	Registration ID	VARCHAR(MAX)	Unique key to identify registrations. Set by Brand X, taken from flat data.
	Registration Date	DATE	Time of device registration. Set by Brand X, taken from flat data.
	Reg_Source_Key	INT	Foreign key to Registration_Sources table, to identify where the device was registered.
	Store_Key	INT	Foreign key to Stores table, to identify where the device was bought.
	Ecomm Flag	BOOL	Boolean ecommerce flag, set to true if the device was purchased online. Taken from flat data.
	Purchase Date	DATETIME	Date that the device was purchased. Reported by Customer, taken from flat data.
	Device Model	VARCHAR(MAX)	Foreign key to Device_Models table, to identify what model the device is.
	Serial Number	VARCHAR(MAX)	Serial number of the device, taken from flat data.

	Customer_Key	INT	Foreign key to Customers table, to identify what customer registered the device.
	Number of Registrations	INT	Number of devices that this customer has registered. Reported by Brand X, taken from flat data.
Messages	Message_Key	INT	Auto Increment key
	Audience	AR(MAX)	Audience that this email was sent to. Taken from flat data.
	Campaign Name	VARCHAR(MAX)	Name of the campaign that this email is a part of. Taken from flat data.
	Version	VARCHAR(MAX)	Creative version from Email Event Data. Taken from flat data
	Subject Line	VARCHAR(MAX)	Email subject line from Email Event Data. Taken from flat data
	Deployment Date	DATETIME	When email has been sent. From Email Event Data. Taken from flat data
Email_Sent	Email_Sent_Key	INT	Auto Increment key
	Email_Key	INT	Foreign Key from Email_Addresses.
	Message_Key	INT	Foreign Key from Messages
	Deployment ID	VARCHAR(MAX)	Unique ID of deployment, set by Brand X. Taken from flat data.
Email_Event_Type	Event_Type_Key	INT	Auto Increment key
	Event Type ID	VARCHAR(MAX)	From Email Event Data. IDs for email types: send, open, click, bounce, unsubscribe. Set by Band X. Taken from flat data.
	Event Type Name	VARCHAR(MAX)	From Email Event Data. Name of the email type(send, open, click). Set by Band X. Taken from flat data.
Email_Link	Link_Key	INT	Auto Increment key
	Message_Key	INT	Foreign Key from Messages.
	Link Name	VARCHAR(MAX)	Link alias. Set from Band X. Taken from flat data.
	URL	VARCHAR(MAX)	The link URL. Set from Brand X. Taken from flat data.
Email_Action	Email_Action_Key	INT	Auto Increment key
	Email_Sent_Key	INT	Foreign Key from Email_Sent table.
	Event_Type_Key	INT	Foreign Key from Email_Event_Type table.
	Event Date	DATETIME	From Email Event Data. Set by Band X. Taken from flat data.

	Link_Key	INT	Foreign Key from Customers table.
--	----------	-----	-----------------------------------