

## **ISTM 637-603: Group Project**

The Data Warehouse Project Report Guidelines and Grading Rubric-

1. **Title Page-** This page contains the project title, course name and number, section number, date, and name of each group member. [1 point]
2. **Table of Contents** [ 1 Point]
3. **Executive Summary [Max 4 points]-** This should not exceed one page. The executive summary will briefly describe the problem statement and your solution to the problem.
4. **Introduction:** In this section include a brief description of the client organization. [**Max 4 points**]- Should NOT exceed one page.
5. **Business Case for Data Warehouse [Max 10 points]-** This should include your analysis of the BI needs of the clients and why these needs cannot be met by their current system. The analysis should include specific examples supporting your assessment/analysis. You will lose points if-
  - You just copy and paste the business intelligence needs of the client organization [- 5 points]
  - You provide a generic write-up of the advantages of Data Warehouse, WITHOUT analyzing whether these advantages apply to the client's case or not. [-5]
6. **Data Warehouse Architecture [Max 10 points]-**
  - The data that will populate the data warehouse is clearly identified (Source and Format) [**Maximum 2 Points**].
  - In the staging component of the architecture, any data transformation required should be mentioned. [**Maximum 2 Points**]
  - The data storage schema (in the data warehouse) is clearly shown and justification provided for choosing that schema. [**Maximum 3 Points**].
  - The reporting component of the architecture should provide client specific details about the BI report users, the devices that they can use to access these reports, the specific BI reports that they can access. Please refer to the client's BI needs to identify the reporting requirements. Use your understanding of the domain to match the BI reports to appropriate user or user category. [**Maximum 3 Points**]
7. **Dimensional Data Design [Max 40 points]-** You will lose points if-
  - Only internal data sources are used [-5]
  - One or more dimensions contain only the most basic details (i.e., the attributes are limited to those available in the internal data source) [-5]
  - The dimension and fact tables are not named as per convention [-2]
  - The fields in the dimension and fact tables are not identified as primary keys, foreign keys, degenerate keys etc. [-3]
  - Facts/Measures used in any fact table are not at the same grain. [-5]
  - Data that should be captured in a fact table is stored in dimension table and vice versa [-5]
  - Your solution to handle slowly changing dimension (SCD) attributes is not reflected in the dimensional design. For example, you decide to address a SCD attribute by using a mini-dimension and this mini-dimension is not shown in the dimensional model. [-5]
8. **Data Extraction, Transformation, and Loading (ETL) [Max 5 points]-** Provide a brief explanation of each SSIS package that you developed, along with the data flow diagram.
9. **BI Reports-** Screenshot of reports that address the primary BI needs of the clients [**Max 10 points**]
10. **Metadata Documentation [Max 10 points]**
11. **Appendices, References, Glossary [Max 3 points]**
12. **Peer Evaluation-** Each student's project grade will be scaled as per his/her peer evaluation score. Please refer to the course syllabus for the details.

# ISTM 637- 603 Data Warehousing Spring 2018

Submitted by

**Group Bikalytics**

Divyesh Batra  
Harikrishnan Subramanyan  
Nupur Sinha  
Tanushree Kothari  
Tsai Mei Chen



## Table of Contents

<b>Executive Summary</b>	3
<b>Introduction</b>	4
<b>Business Case for Data Warehouse</b>	5
<b>Data Warehouse Architecture</b>	8
<b>Dimensional Data Design</b>	Error! Bookmark not defined.
<b>Data Extraction, Transformation, and Loading (ETL)</b>	Error! Bookmark not defined.
<b>BI Reports</b>	Error! Bookmark not defined.
<b>Metadata Documentation</b>	Error! Bookmark not defined.
<b>Appendices</b>	Error! Bookmark not defined.
<b>References</b>	72
<b>Glossary</b>	Error! Bookmark not defined.
<b>Peer Evaluation</b>	73



## Executive Summary

Being the largest charity cycling fundraising campaign, Bike MS has attracted over 70,000 participant to engaged in and has raised \$68 million. [1] However, to raise awareness of multiple sclerosis (MS) and to materialize National Multiple Sclerosis Society's vision of a world free of MS, **Bike MS needs more funding and participants**. Bike MS therefore set a goal for 2018's events: 80,572 riders and 6,489 teams.

To fulfill the goal, we have to possess a thorough understanding of participant and donor behavior. By spotting growth opportunities and determining competitors' weakness, we can develop effective marketing strategy and make better business decisions. Hence, we implement a data warehouse platform using Kimball's bottom-up approach to support the business intelligence solution. We also design a star schema to organize the donation, participation, event, and advertising information.

The results show that energy, financial, and IT industries had the strongest involvement. Among all industries, engineers and managers tend to donate more money. They also show that small events attracts more participants than medium and big events. In addition, the majority of the donors are participants that have no connection with MS. Moreover, the results show ADA – Tour de Cure, Best Buddies, and JDRF are the main competitors of Bike MS. To summarize we suggest MS Bike should hold small or medium events and develop a marketing plan targeting participants and donors meet the following criteria:

1. Participants work in energy, financial, or IT industry.
2. Participants with engineering, executive, or management background.
3. Participants who have no connection with MS or have friends that have MS.
4. Participants in Houston and Philadelphia.
5. Participants using fundraising tools such as Facebook, mobile app, and web page.

1. <http://www.teradatauniversitynetwork.com/Community/Student-Competitions/2018/Data-Challenge/Business-Questions/>

## Introduction

Founded in 1946, National Multiple Sclerosis Society (NMSS) is a non-profit organization based in New York. [1] The organization funds research and programs helping people with multiple sclerosis (MS) and their families.

MS is disabling disease that attacks the central nervous system. Symptoms range widely. In milder cases, there may be numbness in the limbs. Severe cases may involve paralysis or vision loss. [2] Most people with MS are diagnosed between the ages of 20 and 50, with at least two to three times more women than men being diagnosed with the disease. [3] Although there are a number of hypotheses for the causes of MS, such as Epstein-Barr Virus (EBV) and vitamin D deficiencies, the exact cause has not yet been discovered by scientists. [4] According to NMSS, nearly one million people in the United States have MS, and around 10,000 new cases are diagnosed each year. [5]

The mission of NMSS is “people affected by MS can live their best lives as we stop MS in its tracks, restore what has been lost and end MS forever.” To fulfill this mission, in 2016, NMSS devoted more than \$100 million to connect approximately 1 million people affected by MS to the connections, information and resources they need. NMSS also invested \$42 million to support more than 380 new and ongoing research projects around the world. [3] Moreover, NMSS hold fundraising events each year, such as Walk MS, Bike MS, and MuckFest to support people living with MS and to bring people closer to a world free of MS.

1. [https://en.wikipedia.org/wiki/National\\_Multiple\\_Sclerosis\\_Society](https://en.wikipedia.org/wiki/National_Multiple_Sclerosis_Society)
2. <https://www.medicalnewstoday.com/articles/37556.php>
3. [https://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Financials/FY16\\_Annual\\_Progress\\_Report\\_1.pdf?ext=.pdf](https://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Financials/FY16_Annual_Progress_Report_1.pdf?ext=.pdf)
4. <https://multiplesclerosisnewstoday.com/multiple-sclerosis-overview/>
5. <https://multiplesclerosisnewstoday.com/2017/11/20/nearly-1-million-americans-have-multiple-sclerosis-nmss-prevalence-study-finds/>

## Business Case for Data Warehouse

The current business intelligence needs of Bike MS event revolves around two main categories: **Corporate Acquisition** and **Digital/Social Acquisition**. As described on the Bike MS website, the data set is provided in the form of spreadsheets. Based on the data provided, an analysis can be conducted in the following ways:

1. **Using current excel sheets to analyze data**
2. **Using RDBMS approach to analyze data**
3. **Using a data warehouse approach to analyze data**

Based on the analysis done by group BikeAlytics, the approach of designing a data warehouse to analyze data is the best approach pertaining to the case. A detailed explanation of not choosing the other two approaches and selecting a data warehouse approach is presented below:

### **Drawbacks of using current Excel Sheets to Analyze Data:**

**Assertion 1:** With a dataset of around 273.5 MB and an aggregated record size of 4,289,121 rows, the drawback of analyzing data using Excel is that it presents a single view of raw data, all at once which is very difficult to visualize for the senior management of BikeMS.

**Reasoning for single-view of raw data:** Fields like ‘Internal Event Name’, ‘Fiscal Year’, ‘Event ID’ are repeatedly used in various spreadsheets, which makes the data redundant and difficult to comprehend what is important and what is not important to answer business intelligence questions. Thus, generating analytics revolving around multiple spreadsheets with redundant data is not a reliable way to gather insights within a business and predicting future trends to gain competitive advantage.

**Assertion 2:** As the Bike MS team needs to predict future trends based on data gathered during fiscal year 2015-2017, it is imperative to link data from external data sources or other competitors to gain insights into the prevalent market trends. In excel, it becomes a painstaking task to clean data which is not only time-consuming, but also has a potential to introduce errors while cleansing of data.

**Reasoning for Non-integrity of data:** Currently in the data set provided, there are a number of records where there are spelling mistakes such as ‘Constituents’ is misspelled as ‘Constituatu’s’ , ‘Female’ gender is wrongly spelled as ‘Feale’ etc. Thus, with the limited import abilities

spreadsheets offer, processing times can often be very slow and integrity of data is compromised.

**Assertion 3:** As we have to maintain historical records for all the returning participants and corporate teams to Bike MS event, editing the data in same cells lead to permanent loss of historical information and thus results in a potential gap in the analysis.

**Reasoning for Historical Data loss:** Such huge data loss creates problems in data analysis and auditing, making it quite tough to identify trends and evaluate performance without relying on multiple copies of the same spreadsheet from different periods.

### RDBMS vs Data Warehouse Approach

In order to understand the difference between a database and a data warehouse, we must understand that a data warehouse is used for analytical purposes and a database is a repository of data used as a backing data storage for some specific application or set of applications.

For Bike MS event, our main purpose is to analyze data collected during the fiscal years 2013-2017 in order to improve corporate acquisition strategy and digital acquisition strategy.

1. An RDBMS System stores historical information in order to create backups of data, mainly for performance reasons. On the other hand, a data warehouse created for Bike MS allows us to measure facts like '**Total Revenue Collected per Event**' during years 2013-2017 and seek opportunities where further fundraising involvement could have been made for more profits. Owing to denormalized data in data warehouse, simpler joins can be made on tables to reduce the response time for analytical queries. Moreover, insights in specific business area can be answered by specific data marts. For example, in order to check for donations made by various industries, we just need to check donations fact table that has all the measures related to donations.

2. In Data warehouse, we can store additional information from external sources that is bought or captured elsewhere to complement the information that is generated and stored by the internal database system. For example, in order to answer **greatest growth opportunities for new corporate teams**, we have made use of external data sources that yield results for all the corporate organizations that have donated in other events like **Best Buddies, Pedal the Cause, American Cancer Society** etc. and used this additional information to compare donations made in Bike MS events by organizations in the same industry (*energy, engineering, healthcare etc.*). The analysis can be easily shown in the form of reports to senior leadership who can target potential industries to maximize profits in BIKE MS 2018 event.

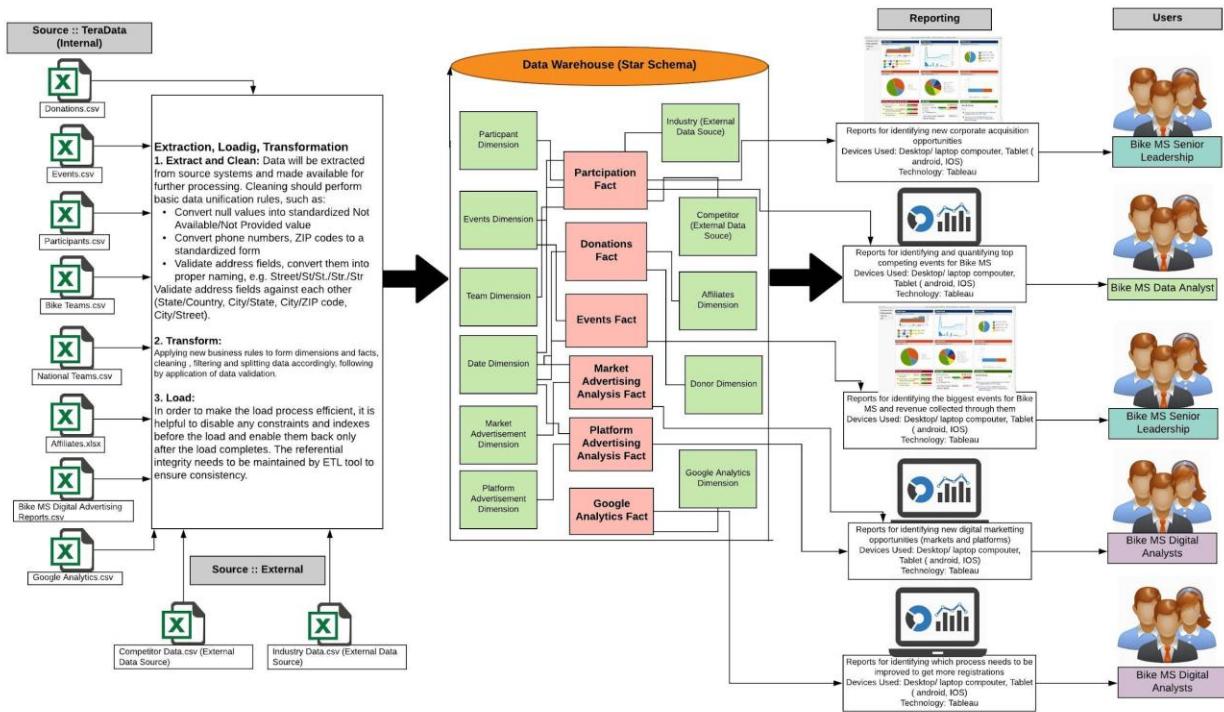
3. Similarly, for digital acquisitions, data from external sources like Google Analytics, Social Media platforms like Facebook, Instagram etc. can be integrated into the data warehouse that provides a quantitative measure to calculate the **cost of clicks-per-ads** (CPA) over the years which clearly answers the **greatest opportunities for digital marketing investments**. For example, based on our analysis, Bike MS should focus on markets like Houston, Philadelphia, NYC, Minneapolis, Denver and Dallas with more focus on video assets and market-specific creatives for Bike MS. Such an analysis is not possible using a traditional RDBMS approach.

4. In order to answer BI questions for Bike MS event, we would have to take into account the grain of information. In data warehouse, we can easily store different versions of the same information, using different structures to compose and access the information. For example, in order to answer BI needs like **corporate acquisitions relevant for biggest events**, we can answer the BI need using 'event-level' data and on the other hand, in order to answer BI needs like **what participant behavior drives fundraising**, we can use databases on 'participant-level'. Hence, the grain of information can significantly vary which supports multiple sets of same data (denormalized data) for data analysis.

Thus, a data warehouse enables us to perform many types of analysis: Descriptive (*what has happened*), Diagnostic (*why it happened*), Predictive (*what will happen*), and Prescriptive (*what to do about it*).

Thus, the future of BikeMS depends on the ability to use the massive amounts of data available to drive better quality at a lower cost. If the senior leadership cannot perform analytics to make sense of the data collected over the years, they will have trouble improving quality and costs, and won't succeed against the prevalent competition in the cycling charity domain. Hence, designing a data warehouse (Online Analytical Processing) using Kimball's approach is of paramount importance.

# Data Warehouse Architecture



## Why Kimball

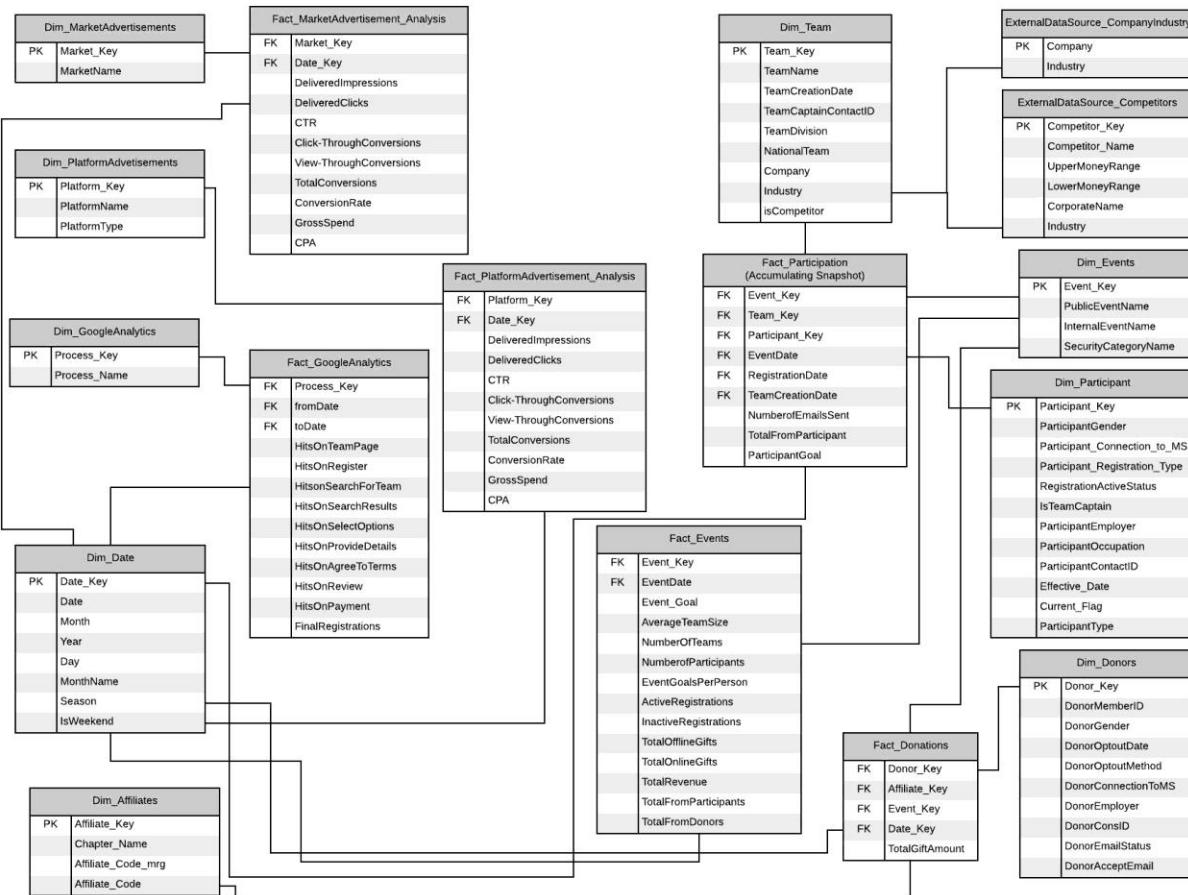
As NMSS wishes to uncover growth opportunities and perform detailed analysis of their company's performance, implementing a data warehouse solution using **Kimball's bottom-up approach** would be most suitable to address their BI needs. It will provide a faster solution, facilitate better coordination and strategic decision-making in the future. A star schema where facts are connected to multiple dimensions would be suitable as it would provide faster access to queries (BI needs).

Kimball's approach would be feasible for NMSS to implement as initially only a single subject area would be developed in phase one. In the subsequent phases, more subject areas would be worked upon, eventually creating the entire data warehouse. This approach will also allow them to develop the skills and techniques required for data warehousing in a much lower risk environment, as they go along the way in an incremental fashion.

Other Benefits of Kimball's approach for NMSS:

- Quick ROI
- Building the data warehouse will take less time
- Initial cost would be low, and each subsequent phase also will cost the same
- Shorter start up time during initial set up as only a single data mart (subject area) would be designed and developed at a time

## Dimensional Data Design



## Data Warehouse Bus Matrix

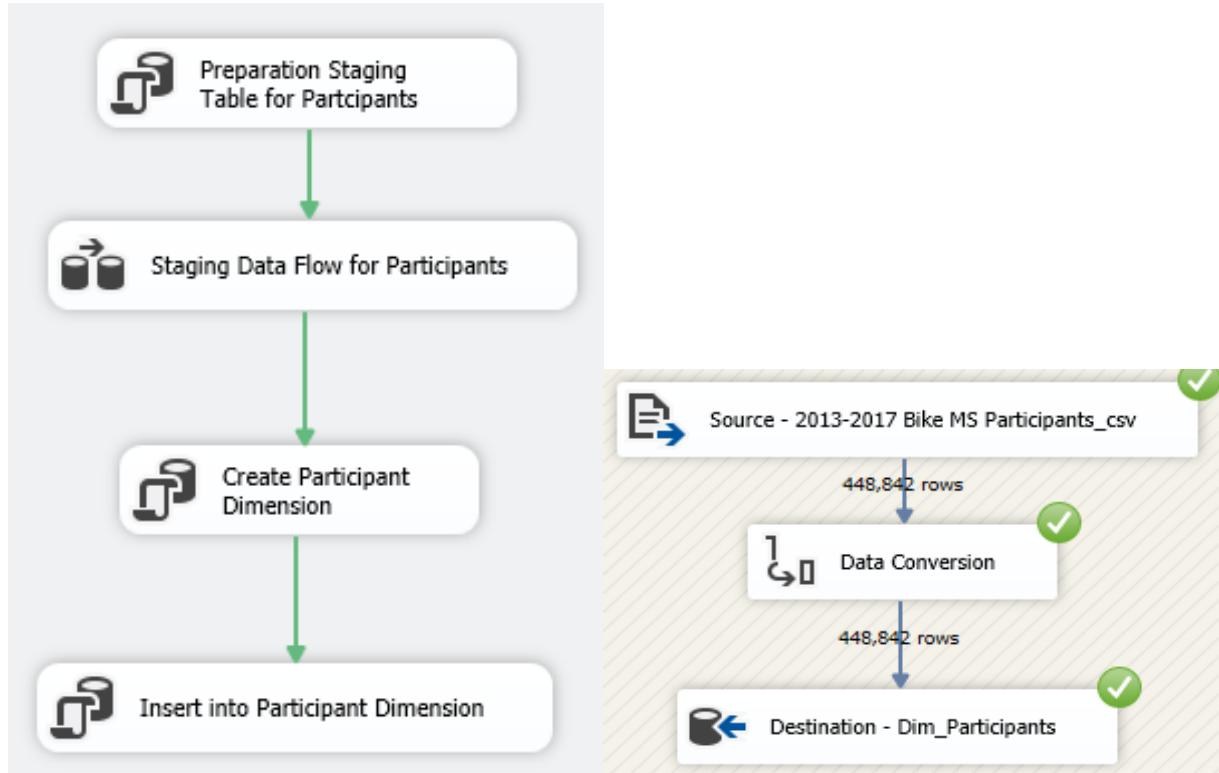
Business Processes	Dimensions								
	Team	Event	Participant	Date	Donors	Affiliates	Google Analytics	Platform Advertisements	Market Advertisements
Events Analysis		X		X					
Donations Analysis		X		X	X	X			

Participants Analysis	x	x	x	x					
Platform Advertisements Analysis				x				x	
Market Advertisements Analysis				x					x
Google Analytics				x			x		

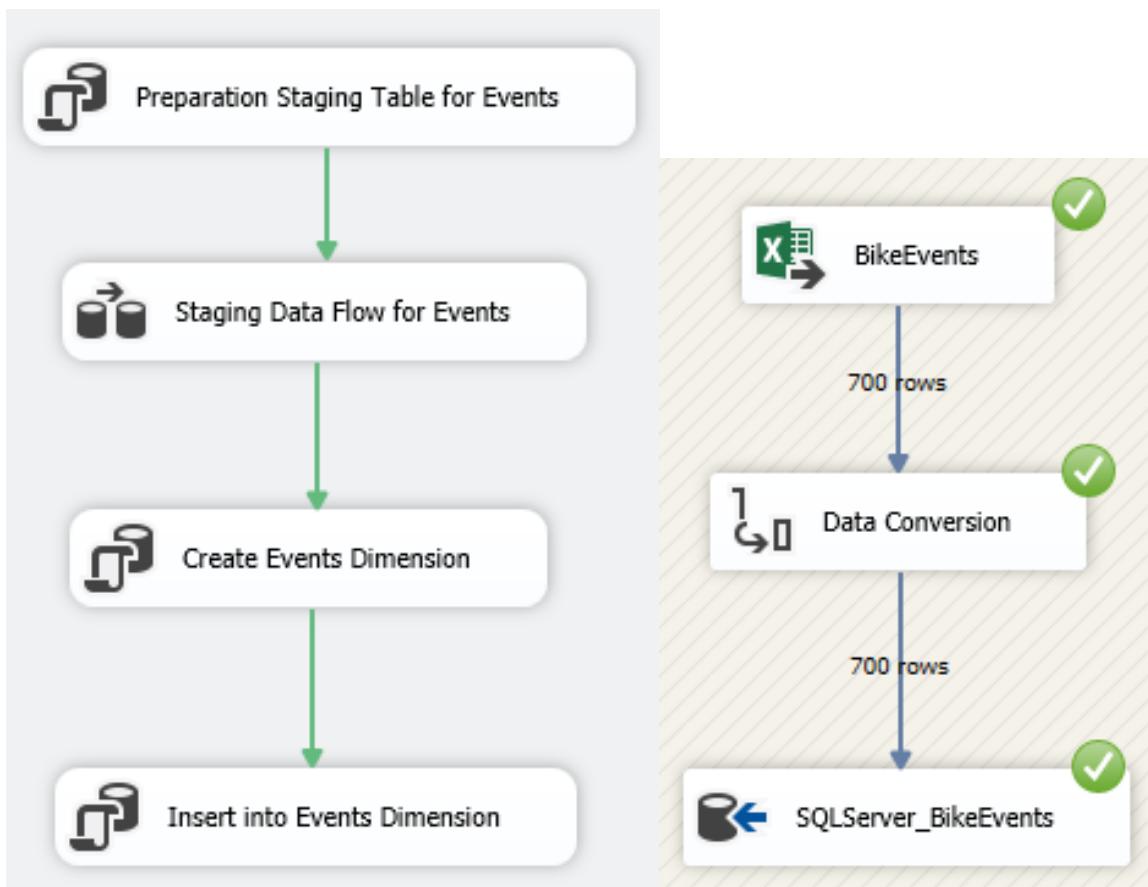
## Data Extraction, Transformation, and Loading (ETL)

The SSIS packages developed, along with the data flow diagram are listed below:

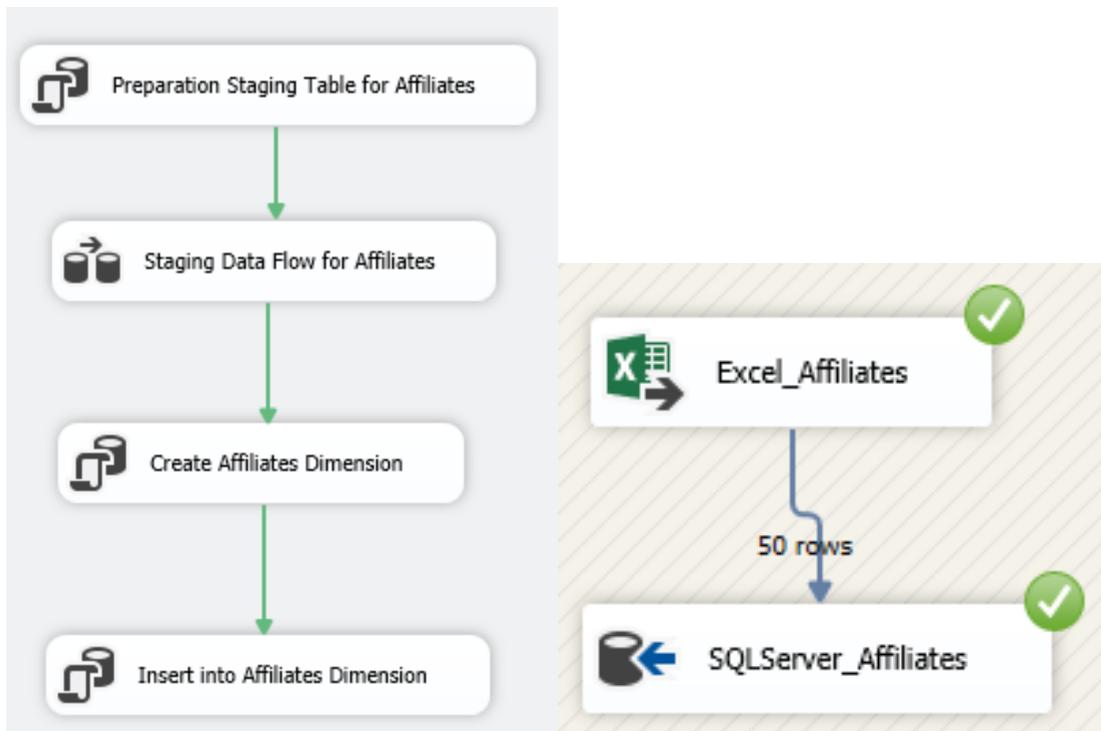
- Participations-** This package is used to load the data from the csv source file '2013-2017 Bike MS Participants' into a staging table. Further, the data is cleaned and filtered to remove bad data, finally transformed using the data conversion transformation to create a participant dimension.



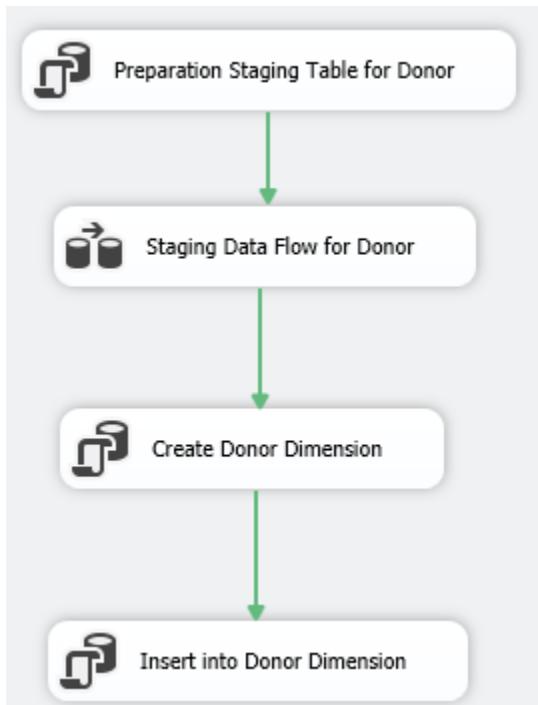
- Events-** This package is used to load the data from the csv source file '2013-2017 Events' into a staging table. Further, the data is cleaned and filtered to remove bad data. As mentioned in the case, the rows having event\_date in the source file as null is test data and therefore is removed during cleaning and is not being considered for analysis. The data is finally transformed using the data conversion transformation to create an event dimension.

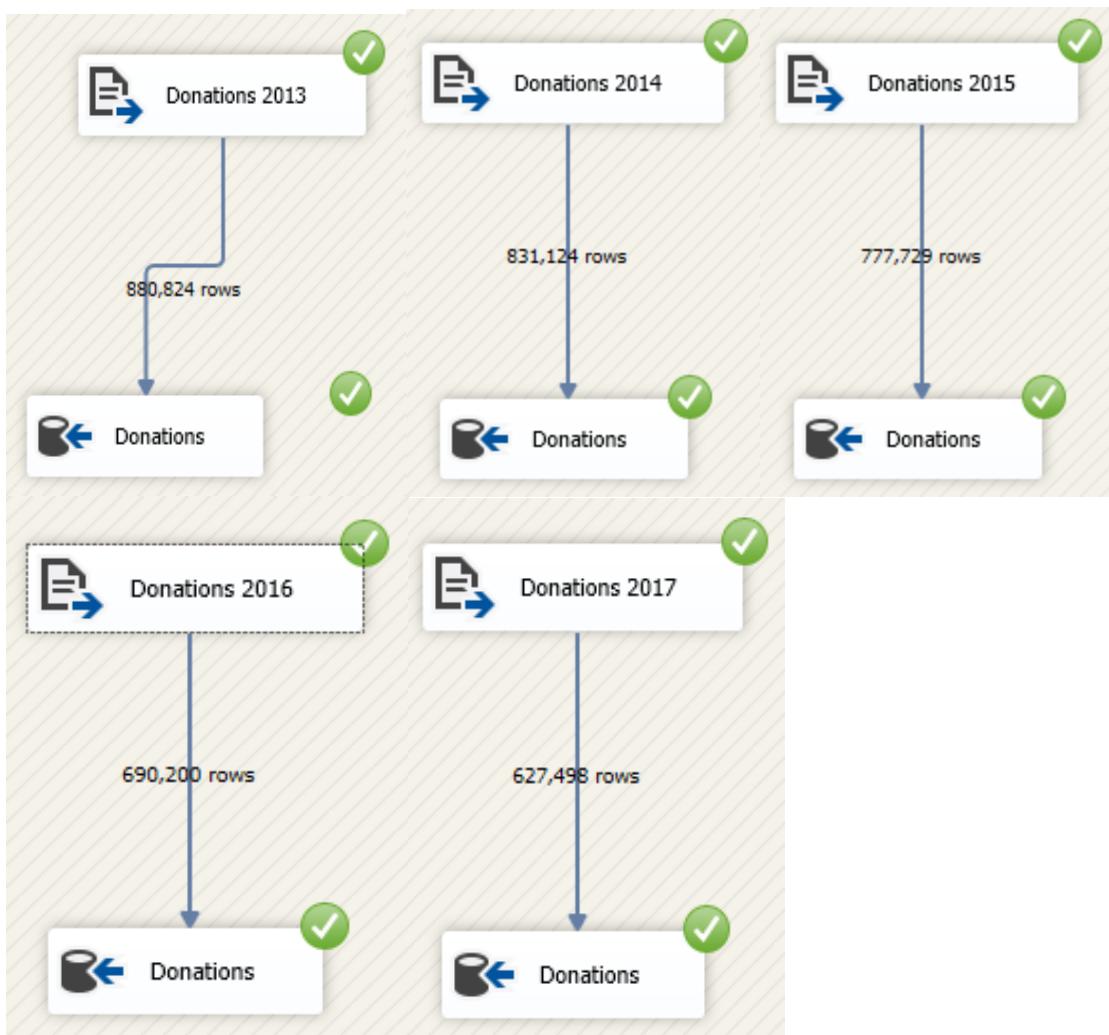


3. **Affiliates**- This package is used to load the data from the csv source file 'Affiliate\_Codes' into a staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally loaded into the target table to create an affiliate dimension.

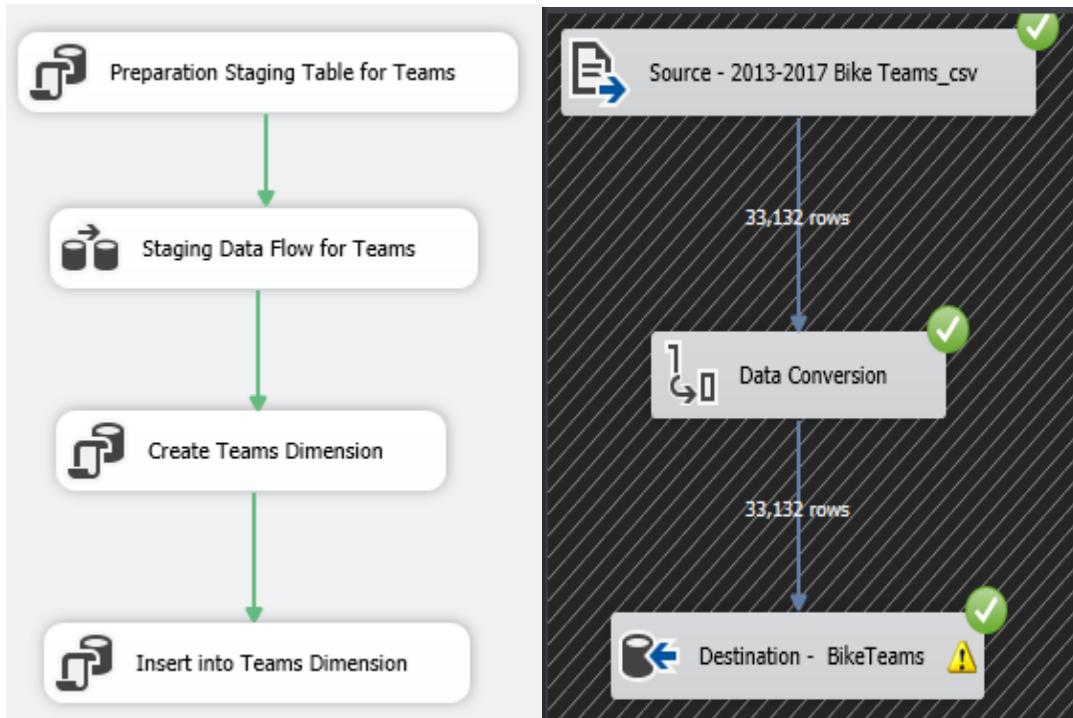


**4. Donations-** This package is used to load the data from five csv source files - 2013 to 2017 Donations into a single staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally loaded into the target table to create a donor dimension.

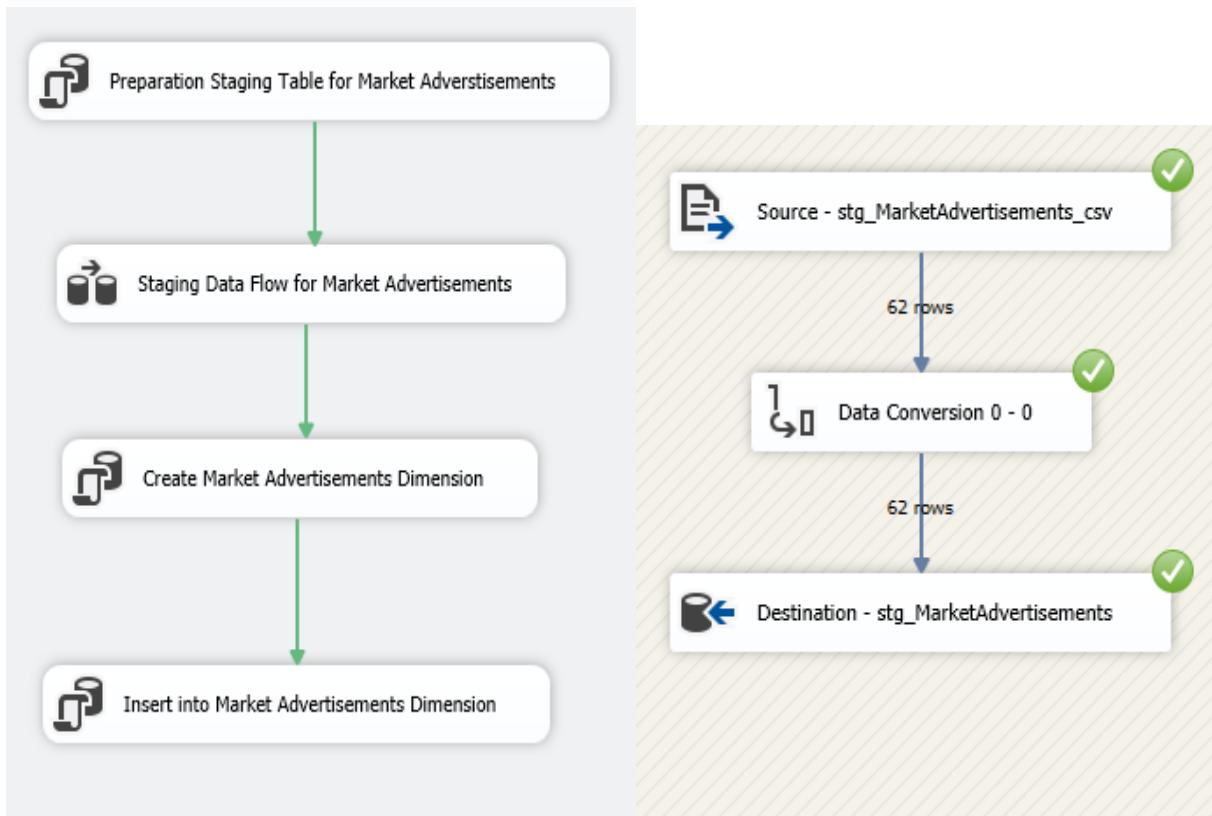




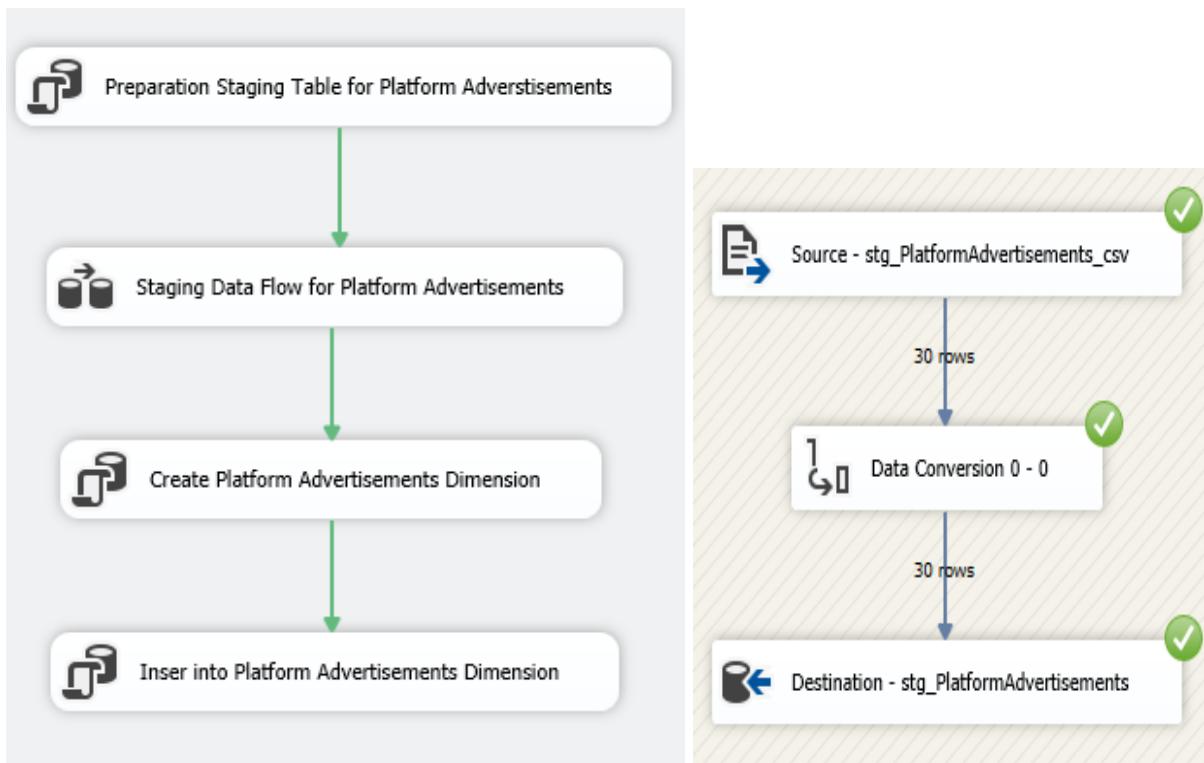
5. **Teams**- This package is used to load the data from csv source file - '2013-2017 Teams' into a staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally transformed using the data conversion transformation to create a team dimension.



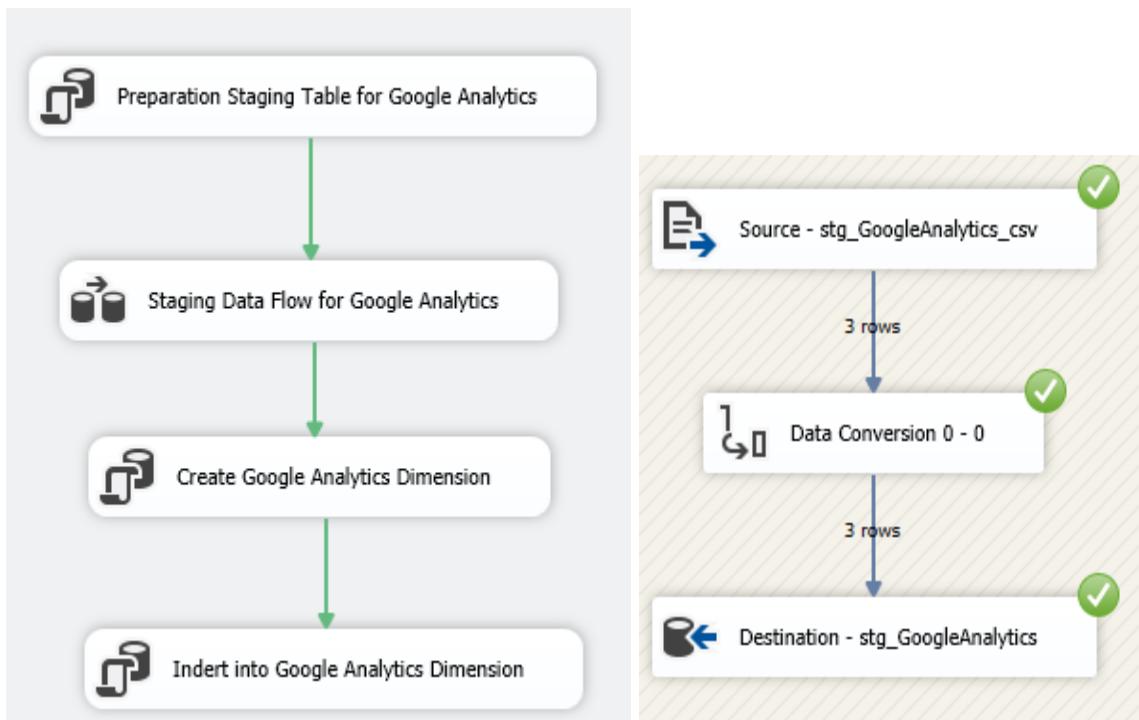
6. **Market Advertisements**: This package is used to load the data from digital marketing source files (Market Advertisements csv source file) into a staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally transformed using the data conversion transformation to create a market advertisements dimension.



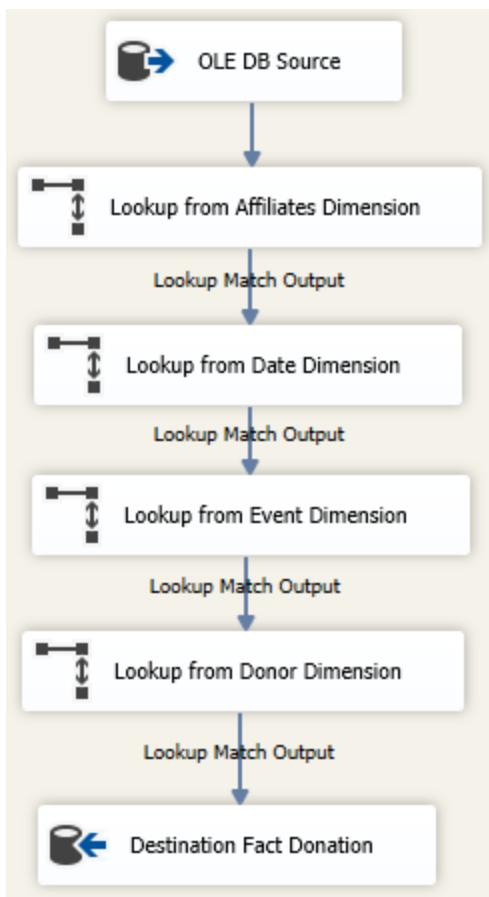
**7. Platform Advertisements:** This package is used to load the data from digital platform source files (Platform Advertisements csv source file) into a staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally transformed using the data conversion transformation to create a platform advertisements dimension.



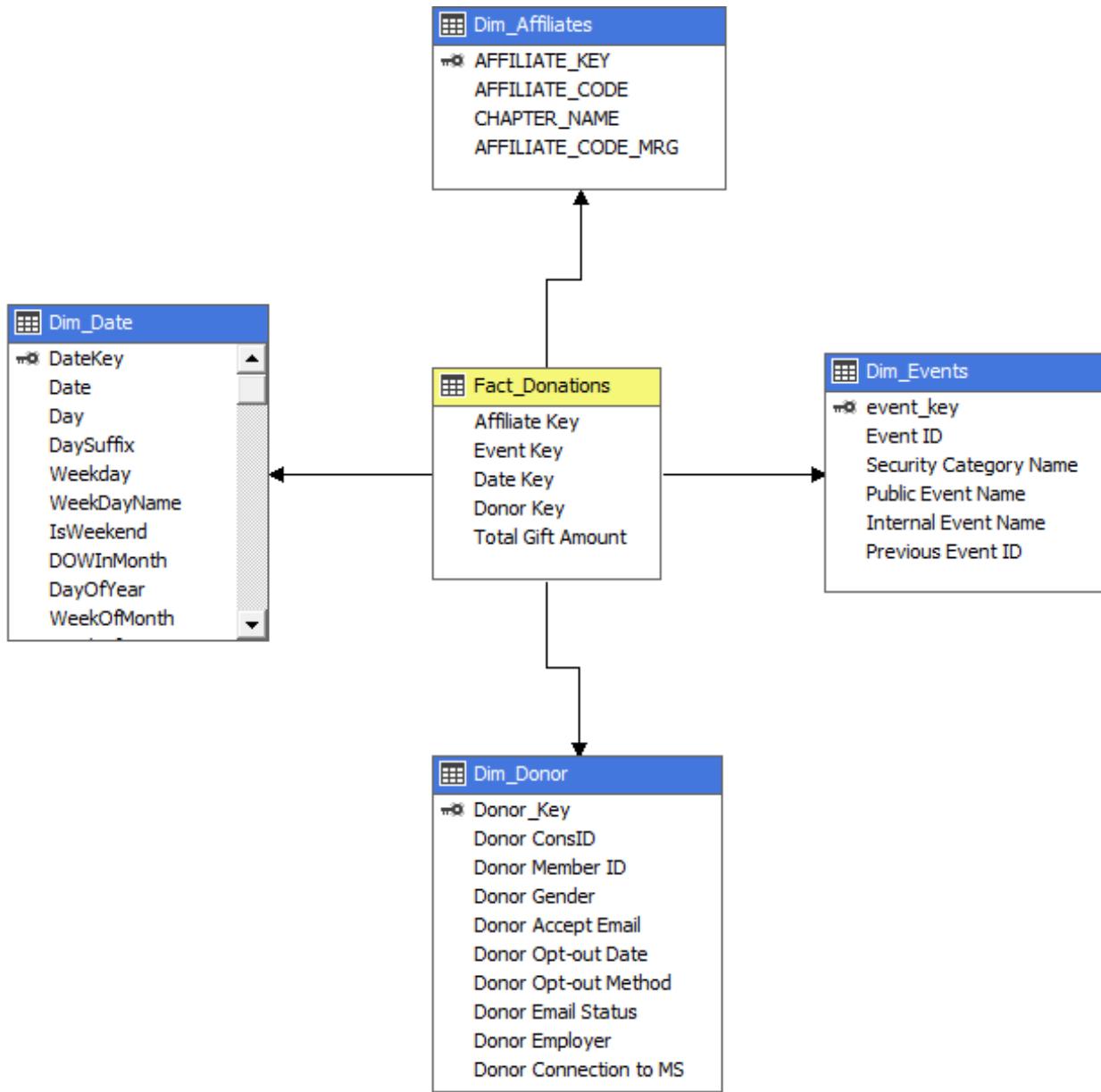
**8. Google Analytics:** This package is used to load the data from Google Analytics source files into a staging table. Further, the data is cleaned and filtered to remove bad data. The data is finally transformed using the data conversion transformation to create a google analytics dimension.



**9. Fact\_Donation:** This package is used to load the data to create a donation fact table by looking up for dimension keys from affiliate dimension, date dimension, event dimension and donor dimension and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the donations.



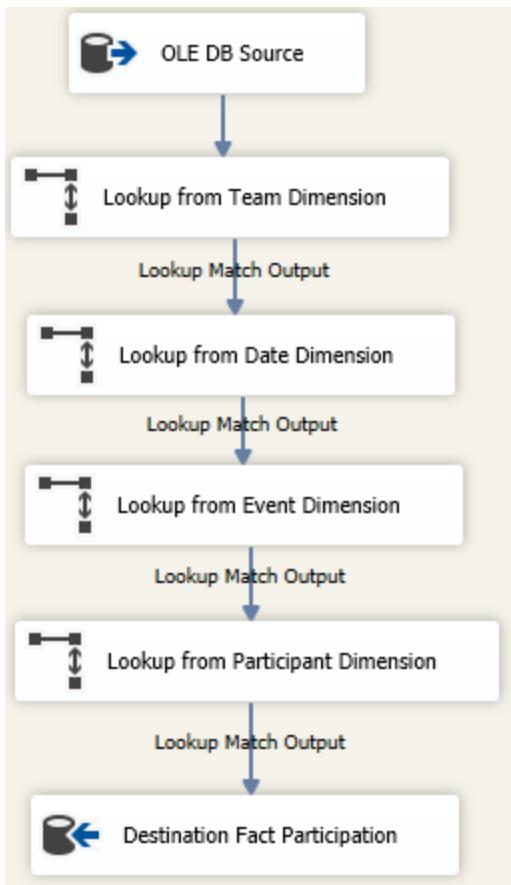
The screenshot shows the Microsoft Visual Studio interface for a multidimensional project named 'MultidimensionalProject1'. The Solution Explorer on the right lists various components: Data Sources (BikeAnalytics.ds), Data Source Views (EventAnalysis.dsv, ParticipationAnalytics.dsv, DonationAnalysis.dsv), Cubes (EventAnalysis.cube, ParticipationAnalytics.cube, DonationAnalysis.cube, Accumulating Snapshot Participation Analysis), Dimensions (DimEvents.dim, DimDate.dim, DimEvents.1dim, DimTeams.dim, DimParticipants.dim, DimAffiliates.dim, DimDate2.dim), and Measures (Measures). The Cube Editor in the center displays the 'Accumulating Snap...\_yisn.cube (Design)' tab, showing a table with columns: AFFILIATE KEY, Date Key, Donor Key, Event Key, Fact Donations Count, and Total Gift Amount. The table contains 25 rows of data. Below the cube editor is the Output window, which shows deployment logs:

```

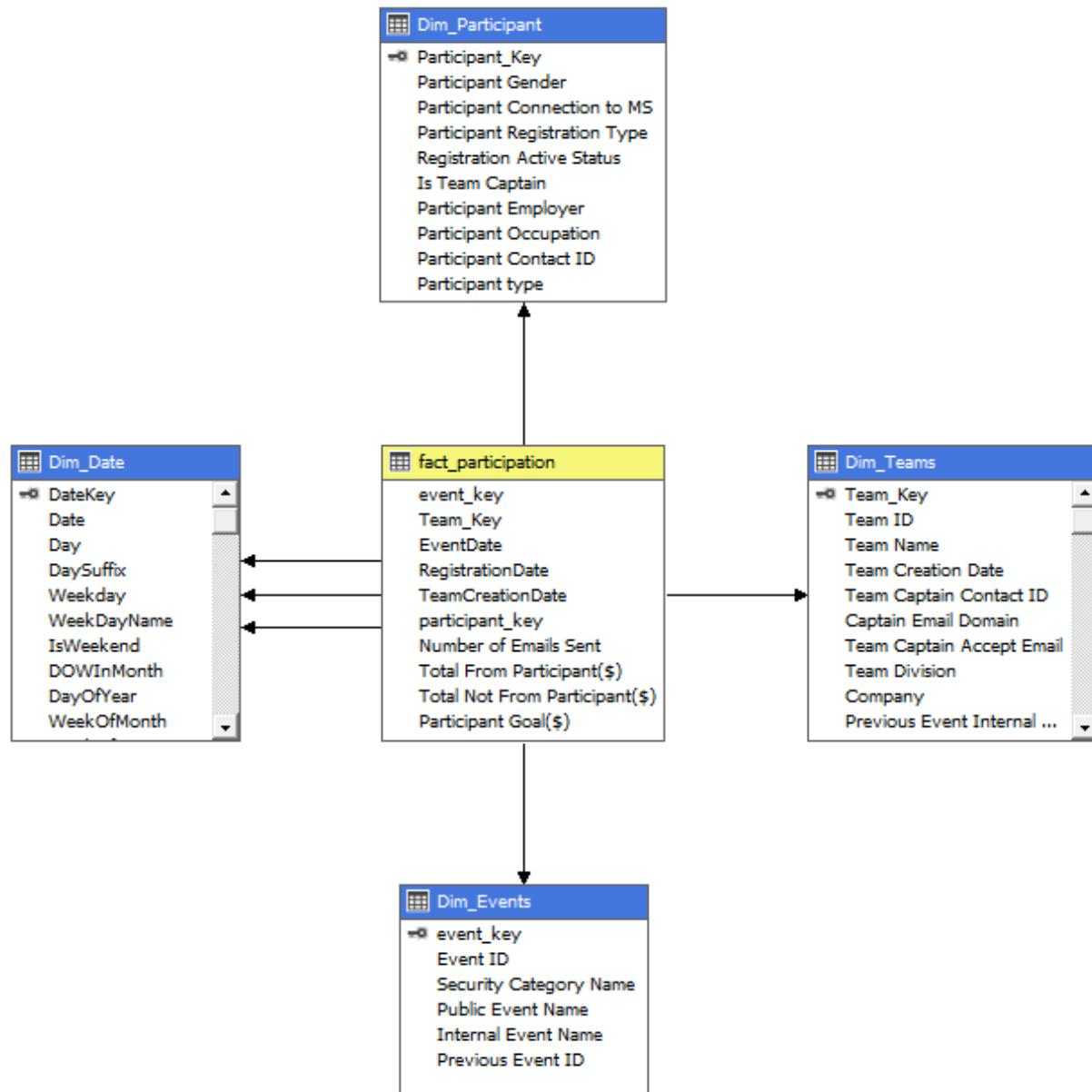
Show output from: Build
Done
Sending deployment script to the server...
Done
Deploy completed... 0 errors, 0 warnings
*****Build: 1 succeeded or up-to-date, 0 failed, 0 skipped *****
*****Deploy: 1 succeeded, 0 failed, 0 skipped *****

```

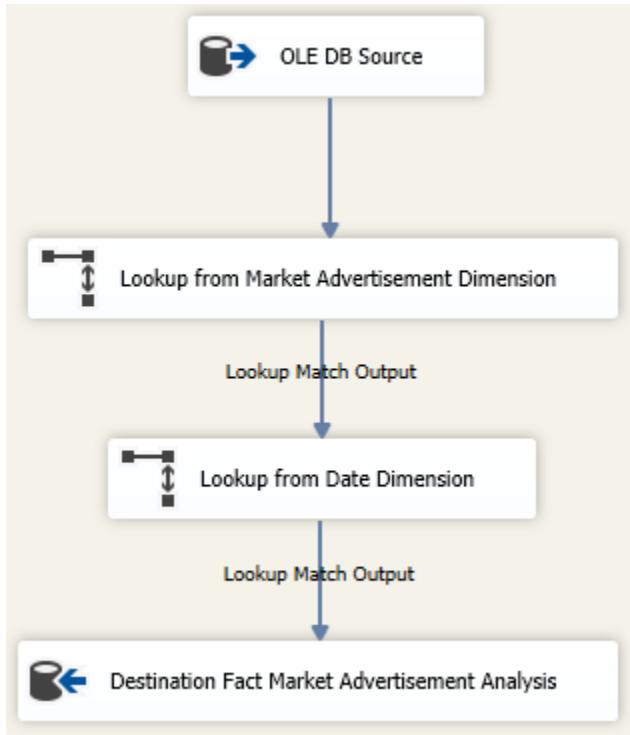
**10. Fact Participation (Accumulating Snapshot)-** This package is used to load the data to create an accumulating snapshot fact table for participation by looking up for dimension keys from team dimension, date dimension, event dimension and participant dimension and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



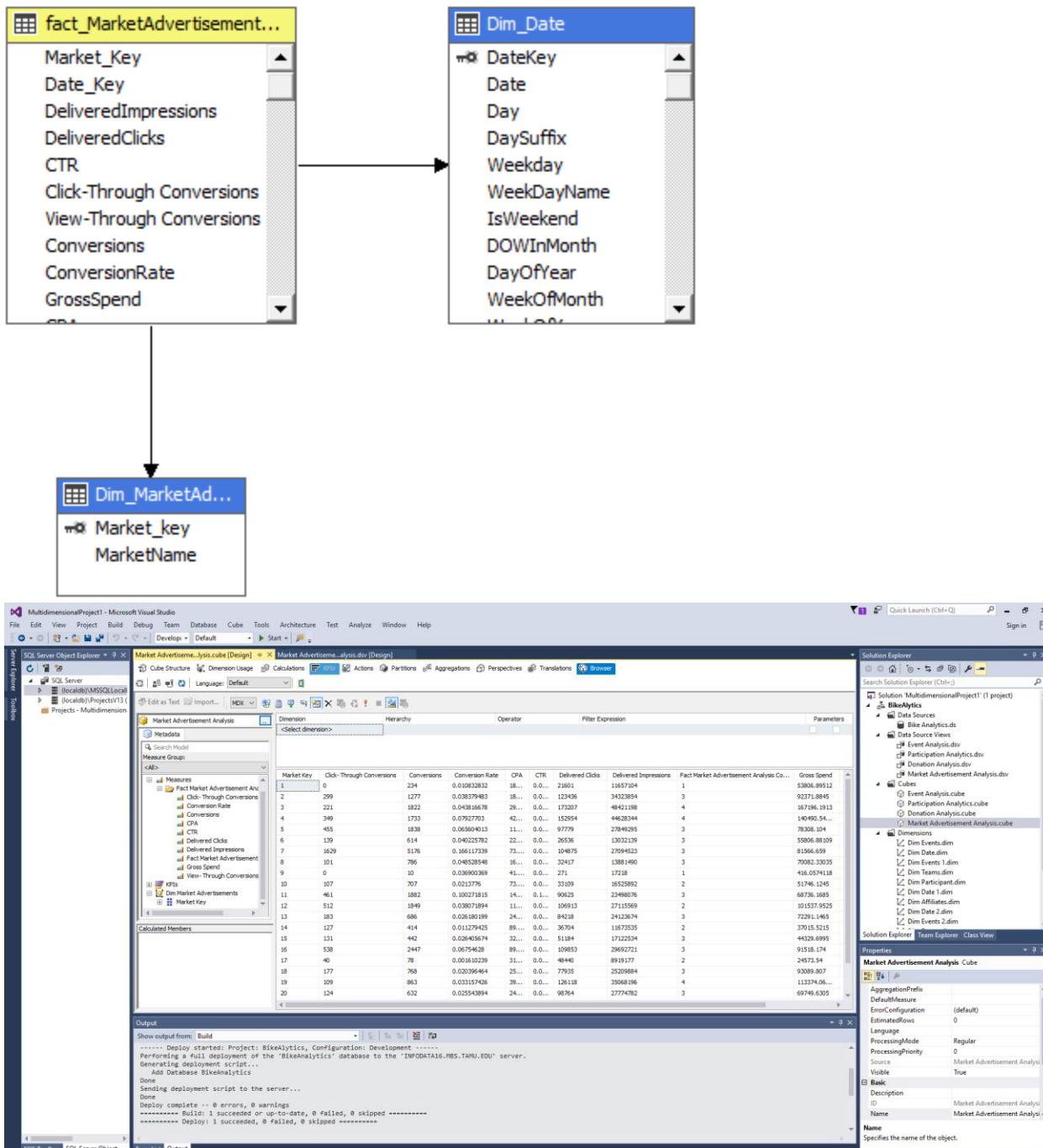
**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the participations in the events.



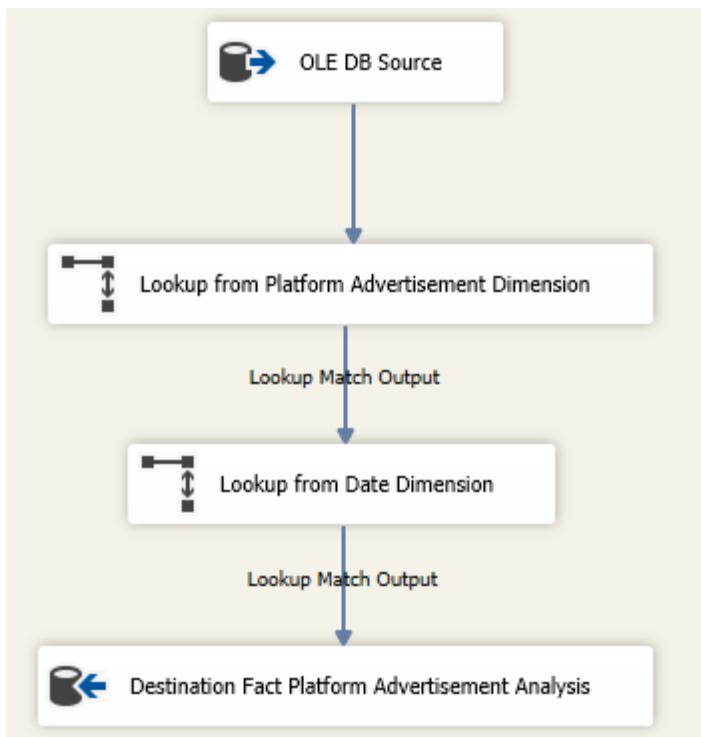
**11. Fact Market Advertisement Analysis-** This package is used to load the data to create a fact table for market advertisement analysis, by looking up for dimension keys from market advertisement dimension and date dimension, and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



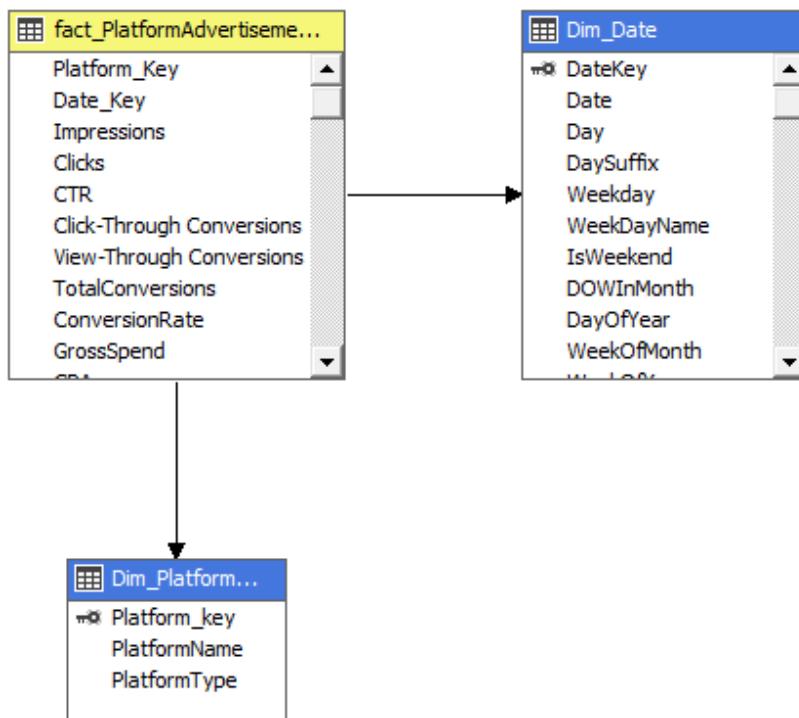
**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the market advertisements.



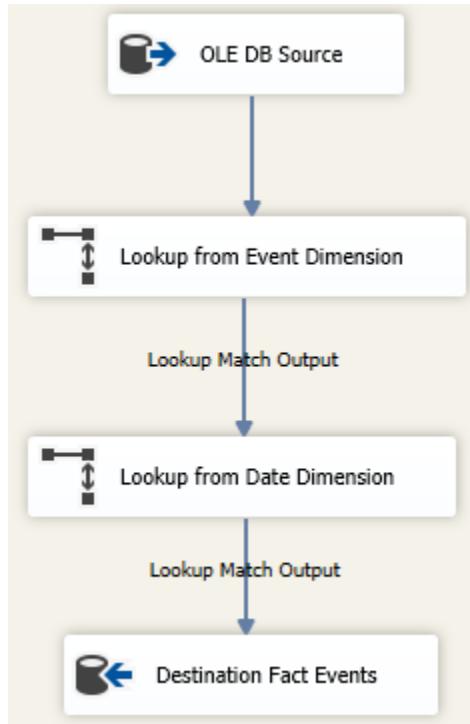
**12. Fact Platform Advertisement Analysis-** This package is used to load the data to create a fact table for platform advertisement analysis, by looking up for dimension keys from platform advertisement dimension and date dimension, and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



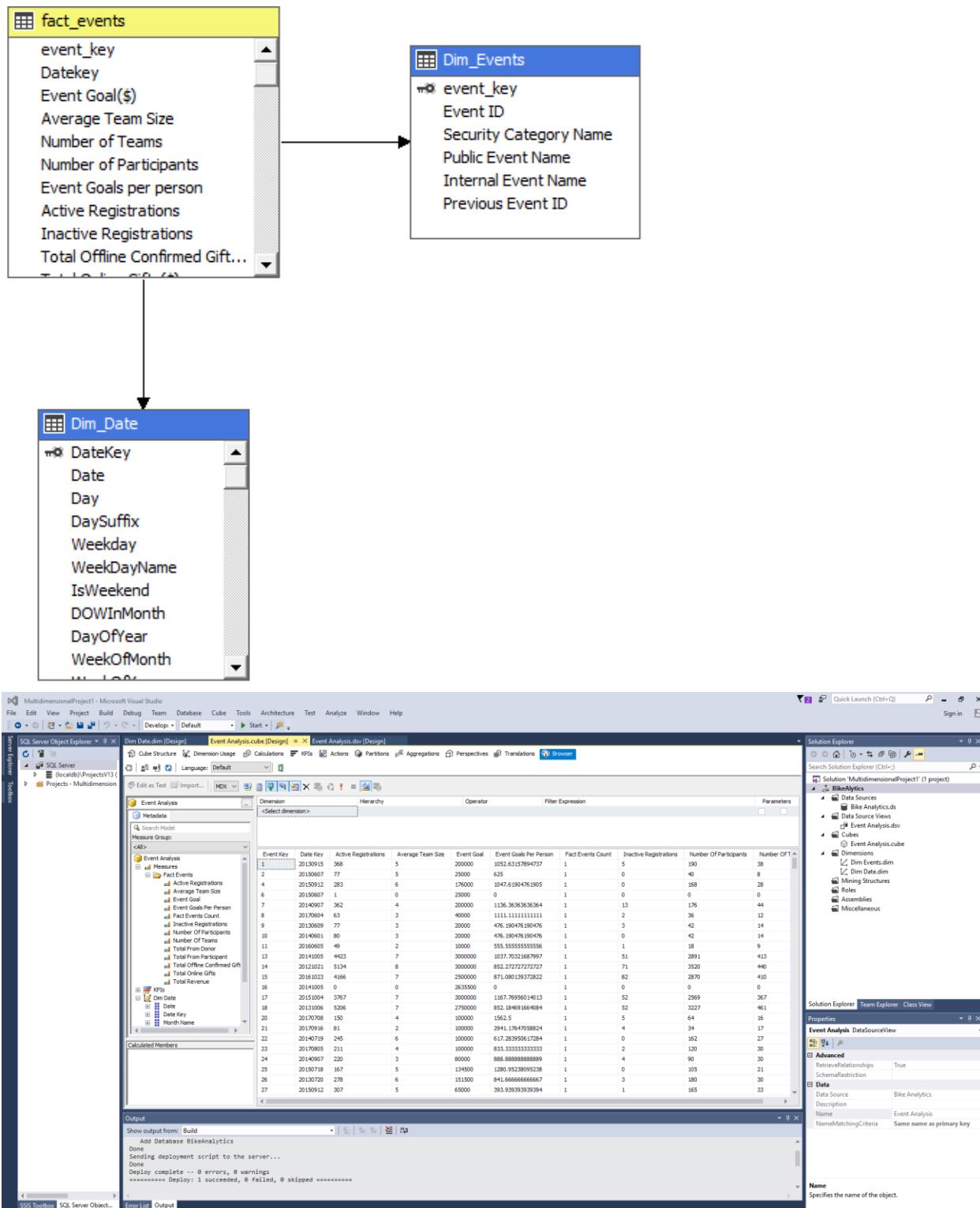
**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the platform advertisements.



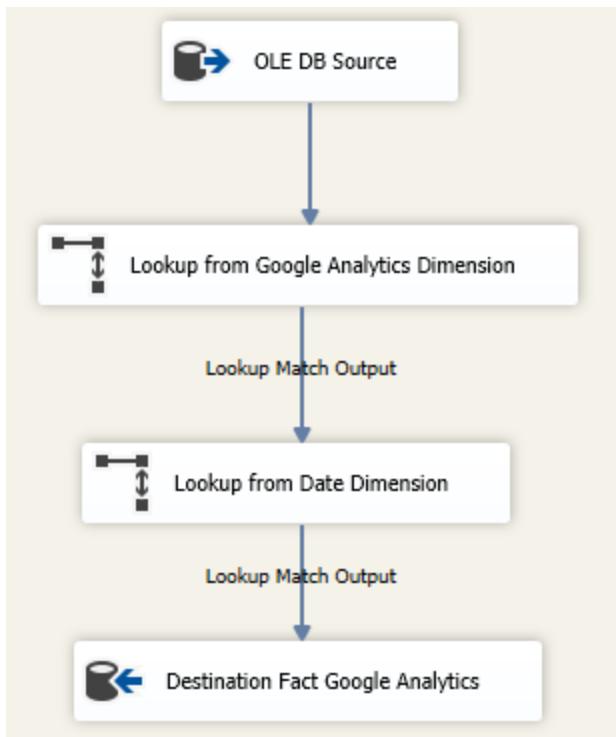
**13. Fact Events:** This package is used to load the data to create an events fact table by looking up for dimension keys from event dimension and date dimension, and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



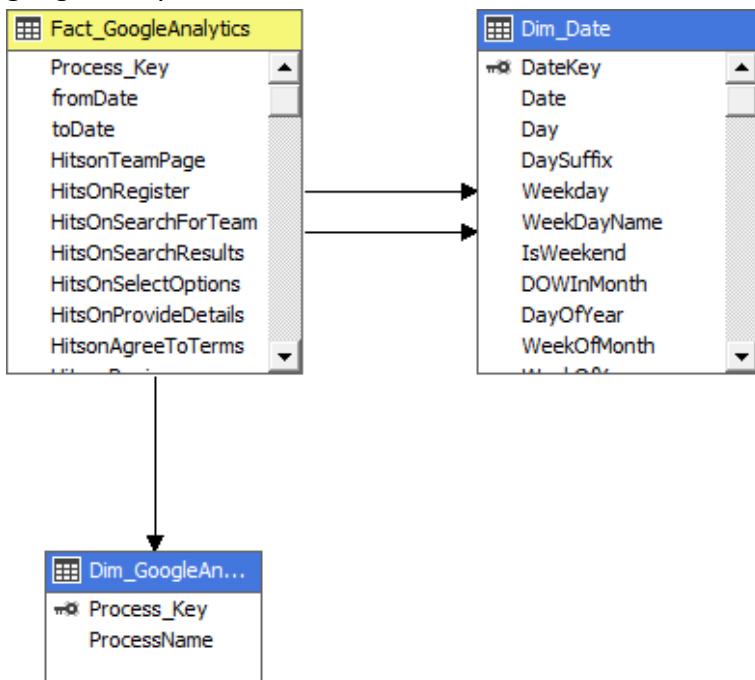
**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the events.



**14. Fact Google Analytics-** This package is used to load the data to create a fact table for google analytics, by looking up for dimension keys from google analytics dimension and date dimension, and measures from corresponding staging tables, aggregated over the mentioned dimension keys.



**Corresponding OLAP Cube-** This cube will be used to answer the BI questions regarding the google analytics.



The screenshot shows the Microsoft Visual Studio interface for a multidimensional project. The main window displays the 'Google Analytics cube [Design]' tab. On the left, the Server Explorer shows a connection to 'SQL Server' and a local database '(localdb)\ProjectV13'. The Solution Explorer on the right lists several projects and their components, including 'BikeAnalytics' with 'Data Sources' like 'Bike Analytics.ds' and 'Data Views' like 'Donation Analysis.dsv'. The Properties pane on the right is set to 'Google Analytics Cube' and shows basic properties such as 'Name' (Google Analytics) and 'Description'.

Process Key	Date Key	Date Key	Fact Google Analytics Count	Final Registrations	Hits On Payment	Hits On Provide Details	Hits On Register	Hits On Search For Team	Hits On Search Results	Hits On Select Options	Hits Agree To Terms	Hits Review	Hits Team Page
1	20160401	20171214	1	16434	22606	22851	49040	0	0	0	34645		
2	20160401	20171214	1	16238	16238	445565	39176						
3	20160401	20171214	1	5713	7401	8502	76540	0					

**Output**

```

Performing a full deployment of the 'BikeAnalytics' database to the 'INFODATA16.MBS.TAMU.EDU' server.
Generating deployment script...
Add Database BikeAnalytics
Done
Sending deployment script to the server...

```

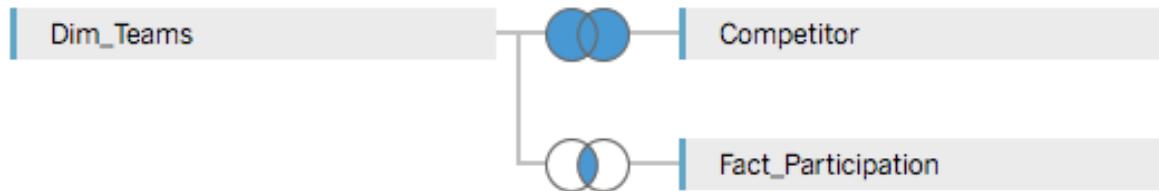
## BI Reports

### ⊖ Dim\_Events+ (Dimensional\_Model)

Connection  
○ Live      ⚡ Extract  
Extract includes all data. 4.



### ⊖ Dim\_Teams+ (Competition-Data)

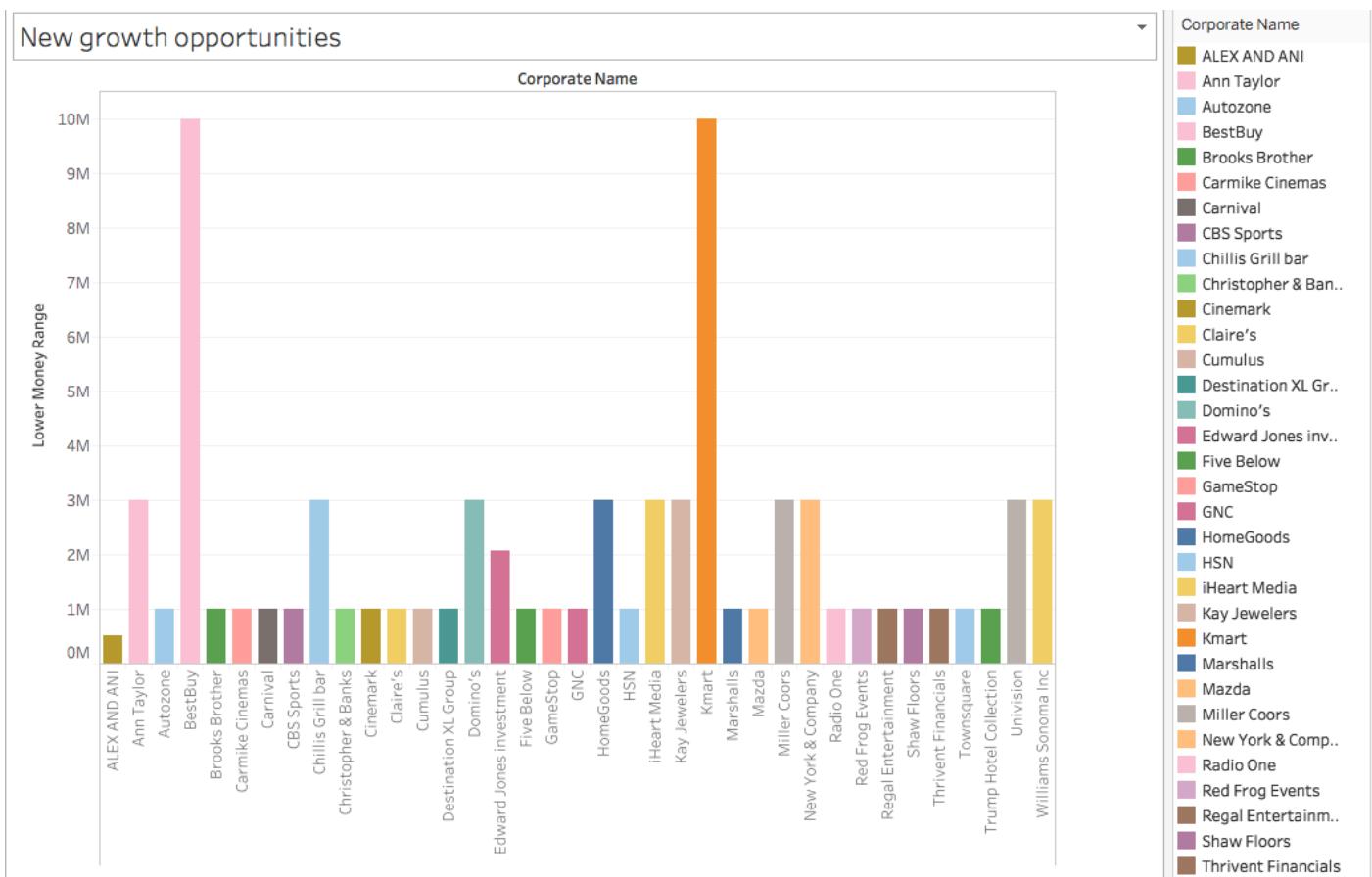


Primary BI needs of NMSS:

## CORPORATE ACQUISITION ANALYSIS REPORTS

BI Q1: What are the greatest growth opportunities for new corporate teams?

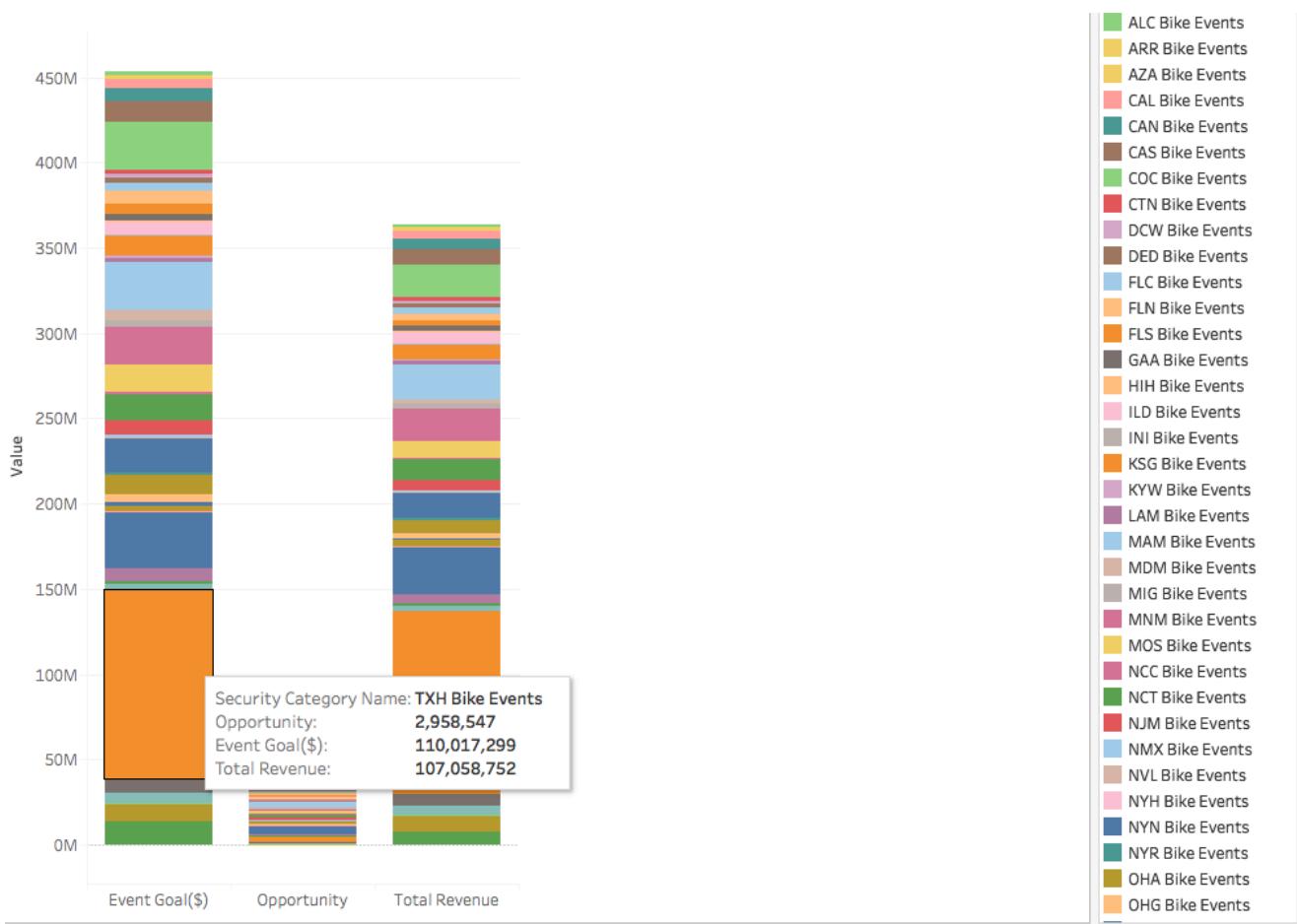
This report shows the donations made by companies in competing events. Companies making high donations could be potential target for NMSS' Bike MS event.



BI Q2: Can we apply those opportunities to specific rides/markets, especially our biggest events?

This report shows the biggest events in terms of opportunity which is a function of event goal and revenue earned. The opportunities identified in the previous section could be applied to these events to meet the event goals.

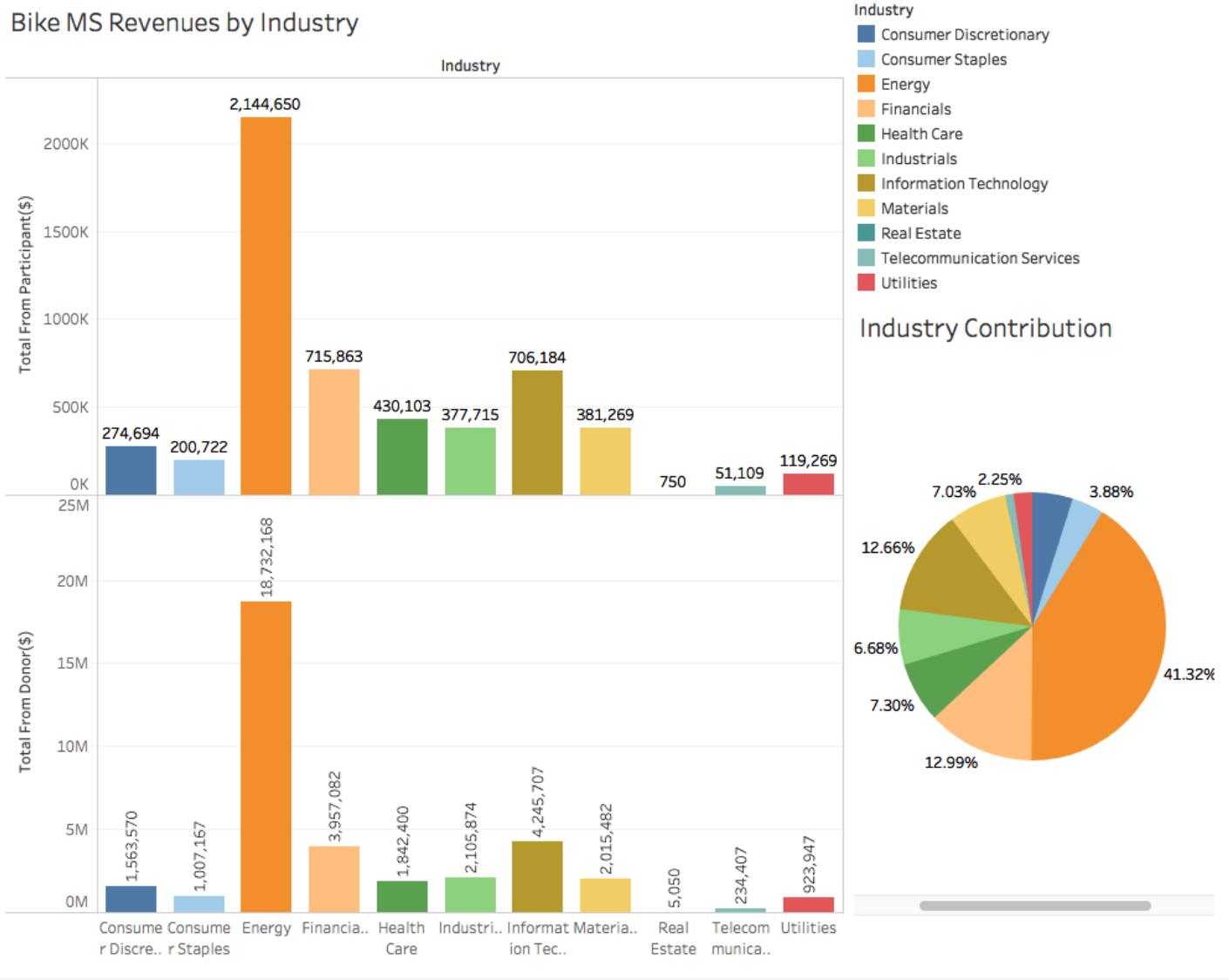
Security Category ..	Event Goal(\$)	Total Revenue	Opportunity=
PAE Bike Events	32,700,000	27,898,942	4,801,058
MAM Bike Events	27,779,000	20,200,321	4,303,679
TXH Bike Events	110,017,299	107,058,752	2,958,547
FLN Bike Events	7,550,000	4,040,605	2,009,395
NCT Bike Events	14,978,500	12,036,022	1,992,478
COC Bike Events	28,425,000	18,888,287	1,641,713
OHA Bike Events	11,246,912	8,140,811	1,606,101
MOS Bike Events	15,970,000	9,532,667	1,537,333
CAN Bike Events	7,687,645	6,200,311	1,487,334
VAR Bike Events	6,140,500	5,195,197	1,274,030
NJM Bike Events	8,500,100	6,317,561	1,182,439
CAL Bike Events	5,404,000	4,493,006	906,994
GAA Bike Events	3,657,000	2,759,683	897,317
MIG Bike Events	3,630,000	2,743,065	861,935
MNM Bike Events	22,211,500	19,283,076	813,425
ILD Bike Events	8,181,000	7,451,262	728,738
PAX Bike Events	7,350,000	5,737,551	712,449
DCW Bike Events	2,512,500	1,826,195	686,305
DED Bike Events	3,050,000	2,337,154	682,846
MDM Bike Events	6,080,000	2,730,456	599,544
OKE Bike Events	1,915,000	1,328,626	586,374
FLC Bike Events	4,021,000	3,455,789	564,211
KSG Bike Events	11,472,381	8,608,649	528,246
OHG Bike Events	4,851,250	2,369,619	491,631
RIR Bike Events	1,595,000	1,115,395	414,605
ALC Bike Events	2,117,000	1,794,141	392,072
WAS Bike Events	10,547,043	10,011,850	346,693
AZA Bike Events	1,903,701	1,557,887	345,814
NYR Bike Events	1,824,500	1,506,401	318,099
TNS Bike Events	3,630,000	3,320,519	309,481
CTN Bike Events	2,458,143	1,975,153	281,244
PAC Bike Events	900,000	341,109	258,891
UTU Bike Events	8,800,000	6,949,615	250,385



BI Q3: What industries have had the strongest involvement in Bike MS in the last five years?

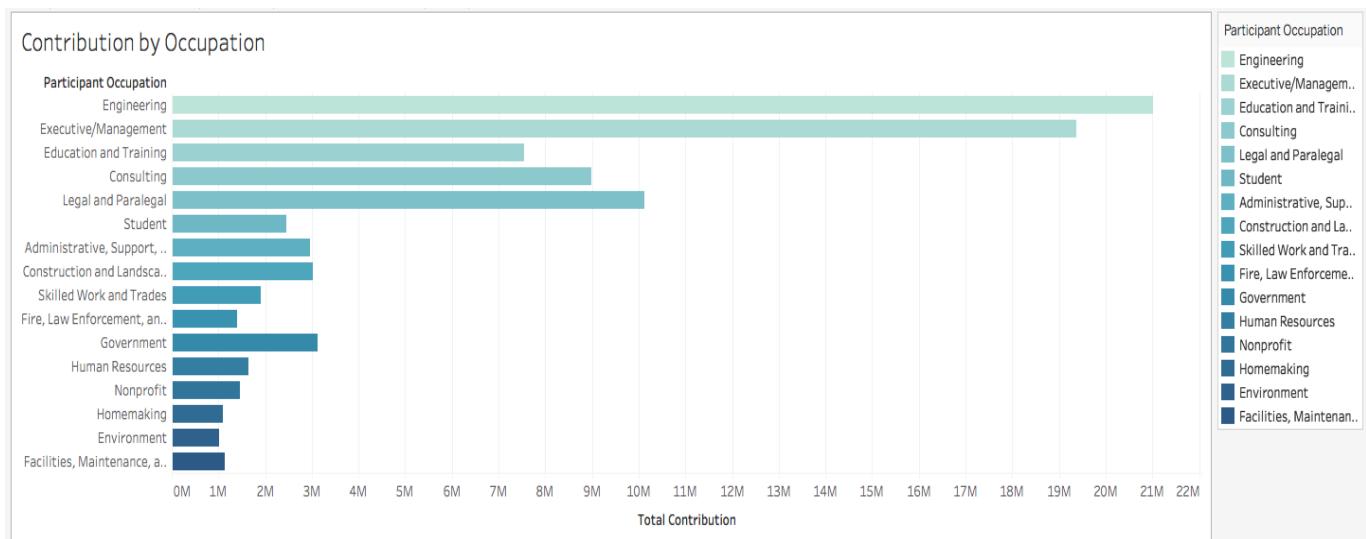
This report shows the industries having the strongest involvement in Bike MS in the last five years.

### Bike MS Revenues by Industry



BI Q4: What occupations were responsible for most of our fundraising?

This report shows the occupations of the participants that were responsible for most of the fundraising.



BI Q5: Can we tie together these industries and occupations to identify gaps/opportunities?

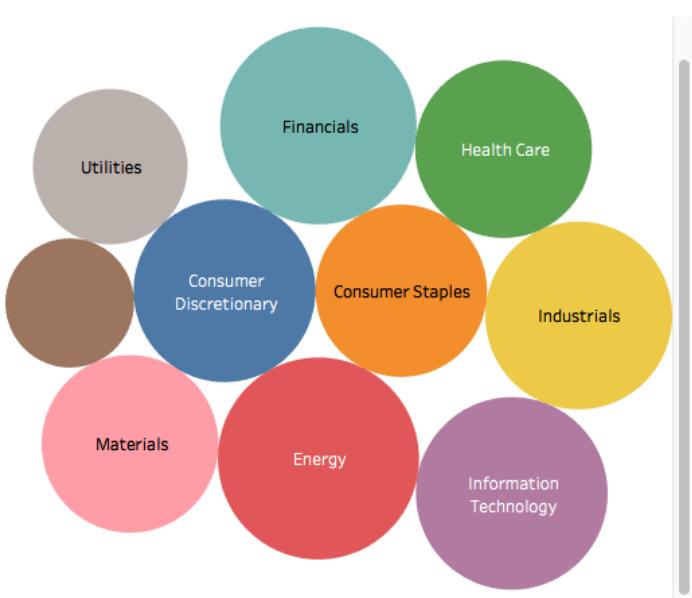
This report shows the number of industries that each occupation targets and vice-versa.

## What Industries and Occupations should Bike MS focus on?

Participant Occupation	
Administrative, Support, and Clerical	10
Construction and Landscaping	10
Consulting	10
Engineering	10
Human Resources	10
Real Estate, Rental, and Leasing	10
Student	10
Architecture	9
Environment	9
Facilities, Maintenance, and Repair	9
Government	9
Insurance	9
Skilled Work and Trades	9
Military	8
Nonprofit	8
Homemaking	7
Stock Broker/Investment Advisor	6
Clergy	5
Social Work	5
Publishing	3
Photography	2
Manager	1
Retired	1
Technical Account Manager	1
Vice President- Wealth Management	1

Industry	
Consumer Discretionary	18
Consumer Staples	16
Energy	22
Financials	21
Health Care	17
Industrials	19
Information Technology	20
Materials	17
Telecommunication Servi..	9
Utilities	13

Industry
Consumer Discretionary
Consumer Staples
Energy
Financials
Health Care
Industrials
Information Technology
Materials
Telecommunication Services
Utilities



BI Q6: What is the common denominator for our top performing corporate teams? (Is it industry, culture, executive involvement, connection to MS, other?)

This report shows the common denominator for the top performing corporate teams as the connection to MS. Having a connection to MS significantly drives high performing corporate teams.

### Common Denominator for high contributions

Participation Connection	Industry										
	Energy	Informati on Techn..	Financials	Industri..	Health Care	Consumer Discretio..	Consumer Materials	Staples	Utilities	Telecomm unicatio..	Real Estate
Caregiver	8	7	2	1	2		8	2			
Family	2,750	1,149	1,156	755	660	700	446	347	257	75	1
Friend	5,239	2,215	2,095	1,609	1,506	1,127	1,191	840	374	196	3
No connection	5,694	2,295	2,172	1,555	1,060	1,153	993	765	470	135	13
Other	947	318	432	358	298	147	175	123	41	37	6
Personally infected	156	165	148	76	80	84	30	48	27	13	
SUM(Number of Records)											
										1	5,694

BI Q7: Can we quantify the effect competing events are having in our top markets?

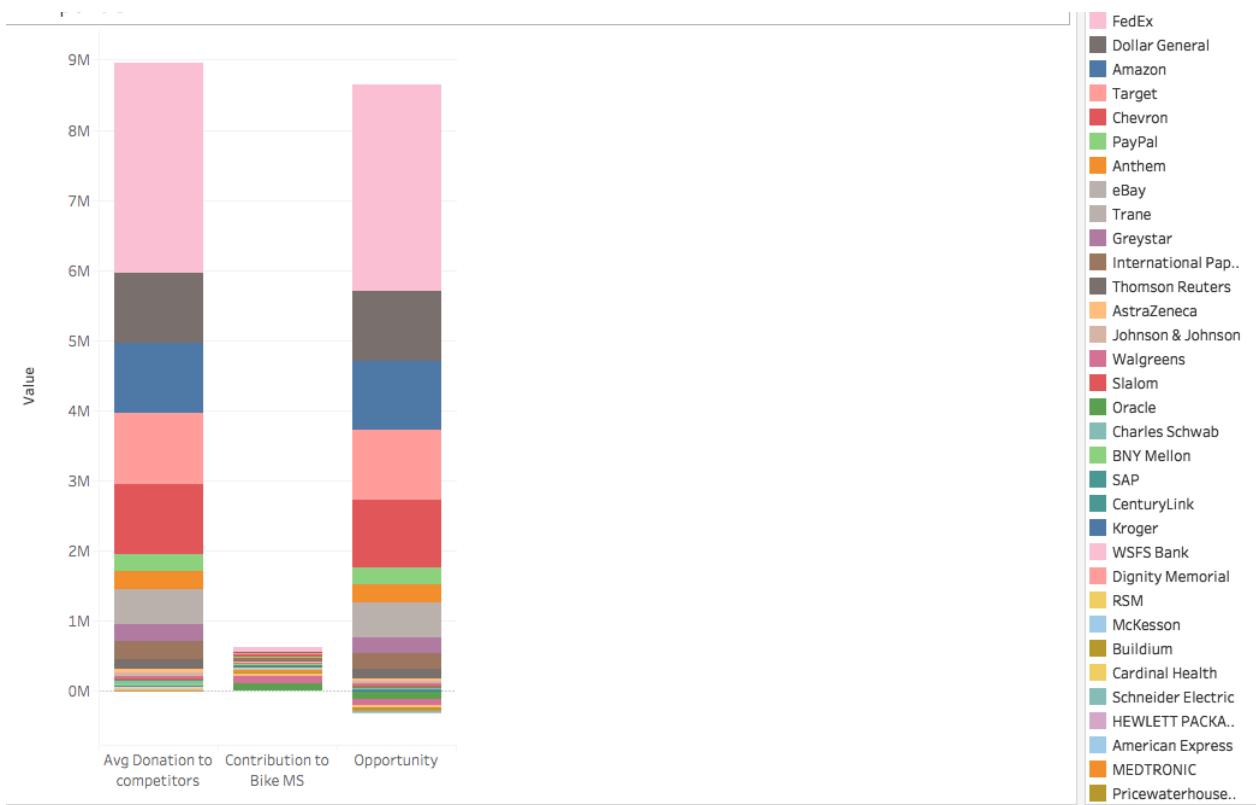
The following reports shows the effect of the below competitors that are affecting the top markets of Bike MS:

1. ADA – Tour de Cure
2. Best Buddies
3. JDRF
4. PedalTheCause
5. St. Jude Children's Research Hospital

Corporate Name	ADA - Tour de Cure	BestBuddies	JDRF	PedalTheCause	St. Jude Children's R..
Amazon					53,000,000
American Express	4,991,378	1,580,000			
Anthem					500,000
AstraZeneca	129,315				
AT&T	253,470	2,130,000			
BNY Mellon		3,250,000			
Buildium		285,000			
Cardinal Health	477,090				
CenturyLink	148,440				
Charles Schwab	6,235,504				
Chevron					223,000,000
Dignity Memorial	3,090				
Dollar General					23,000,000
eBay					5,750,000
Ernst & Young	2,327,300				
FedEx					2,589,000,000
Greystar					28,000,000
Hewlett Packard	11,900				
International Paper					58,250,000
Johnson & Johnson	11,034,540				
Kroger	921,352				
McKesson	660				
Medtronic		1,305,680			
Merck	29,625				
Oracle	74,585				
Paypal					500,000
PriceWaterhouseCoopers ..		875,000			
RSM	12,600				
SAP	179,265				

The following two report shows the potential corporate acquisitions opportunity for Bike MS.

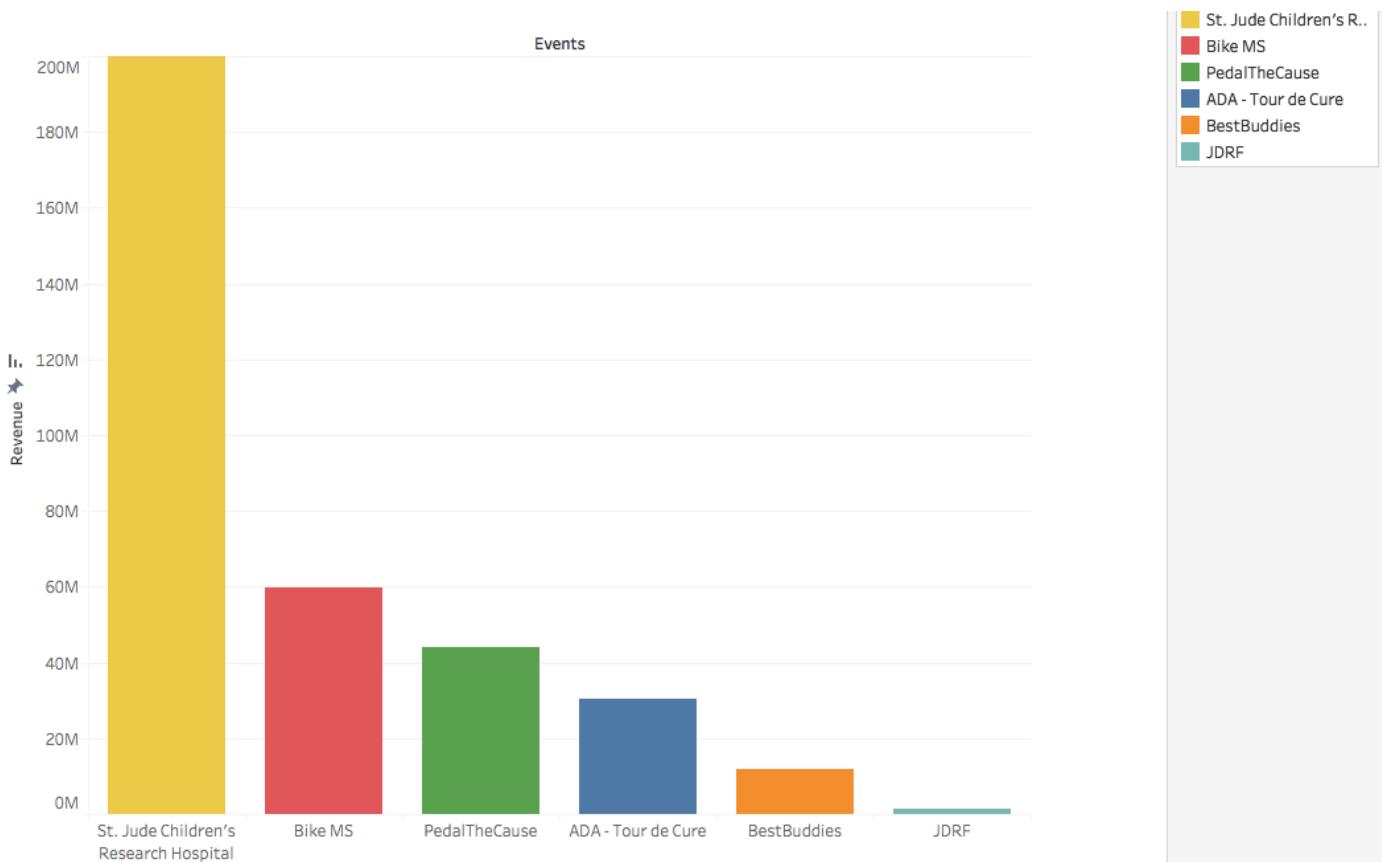
Company	Avg Donation to competitors	Contribution to Bike MS	Opportunity=
FedEx	3,000,000	56,407	2,943,593
Dollar General	1,000,000	6,635	993,365
Amazon	1,000,000	10,200	989,800
Target	1,000,000	17,248	982,752
Chevron	1,000,000	31,046	968,954
PayPal	250,000	55	249,945
Anthem	250,000	215	249,785
eBay	250,000	250	249,750
Trane	250,000	1,390	248,610
Greystar	250,000	19,903	230,097
International Paper	250,000	30,416	219,584
Thomson Reuters	147,166	16,242	130,924
AstraZeneca	43,105	150	42,955
Johnson & Johnson	55,730	21,575	34,155
Walgreens	34,754	7,457	27,297
Slalom	17,003	65	16,938
Oracle	14,917	40	14,877
Charles Schwab	35,429	21,308	14,121
BNY Mellon	25,000	12,779	12,221
SAP	9,435	3,326	6,109
CenturyLink	7,422	2,150	5,272
Kroger	11,236	7,353	3,883
WSFS Bank	5,000	1,160	3,840
Dignity Memorial	3,090	0	3,090
RSM	1,400	45	1,355
McKesson	110	1,400	-1,290
Buildium	5,000	7,189	-2,189
Cardinal Health	5,022	7,399	-2,377
Schneider Electric	100	6,510	-6,410
HEWLETT PACKARD	350	7,278	-6,928



This report shows the donations made by corporate teams to all the competing events as well as Bike MS.

Corporate Name	ADA – Tour de Cure	BestBuddies	Competitor Name	JDRF	PedalTheCause	St. Jude Children's R..	
Amazon						53,000,000	10,200
American Express	4,991,378	1,580,000					28,960
Anthem						500,000	215
AstraZeneca	129,315						150
AT&T	253,470	2,130,000					114,954
BNY Mellon		3,250,000					12,779
Buildium		285,000					7,189
Cardinal Health	477,090						7,399
CenturyLink	148,440						2,150
Charles Schwab	6,235,504						21,308
Chevron						223,000,000	31,046
Dignity Memorial	3,090						0
Dollar General						23,000,000	6,635
eBay						5,750,000	250
Ernst & Young	2,327,300						98,261
FedEx						2,589,000,000	56,407
Greystar						28,000,000	19,903
Hewlett Packard	11,900						7,278
International Paper						58,250,000	30,416
Johnson & Johnson	11,034,540						21,575
Kroger	921,352						7,353
McKesson	660						1,400
Medtronic			1,305,680				20,340
Merck	29,625						33,950
Oracle	74,585						40
Paypal					500,000		55
PriceWaterhouseCoopers ..		875,000					24,635
RSM	12,600						45
SAP	179,265						3,326
Schneider Electric	6,800						6,510
Slalom	17,003						65
Target						231,000,000	17,248
Thomson Reuters					41,353,646		16,242

This report shows the rank of each competing event in terms of revenue.



## DIGITAL/SOCIAL ACQUISITION REPORTS

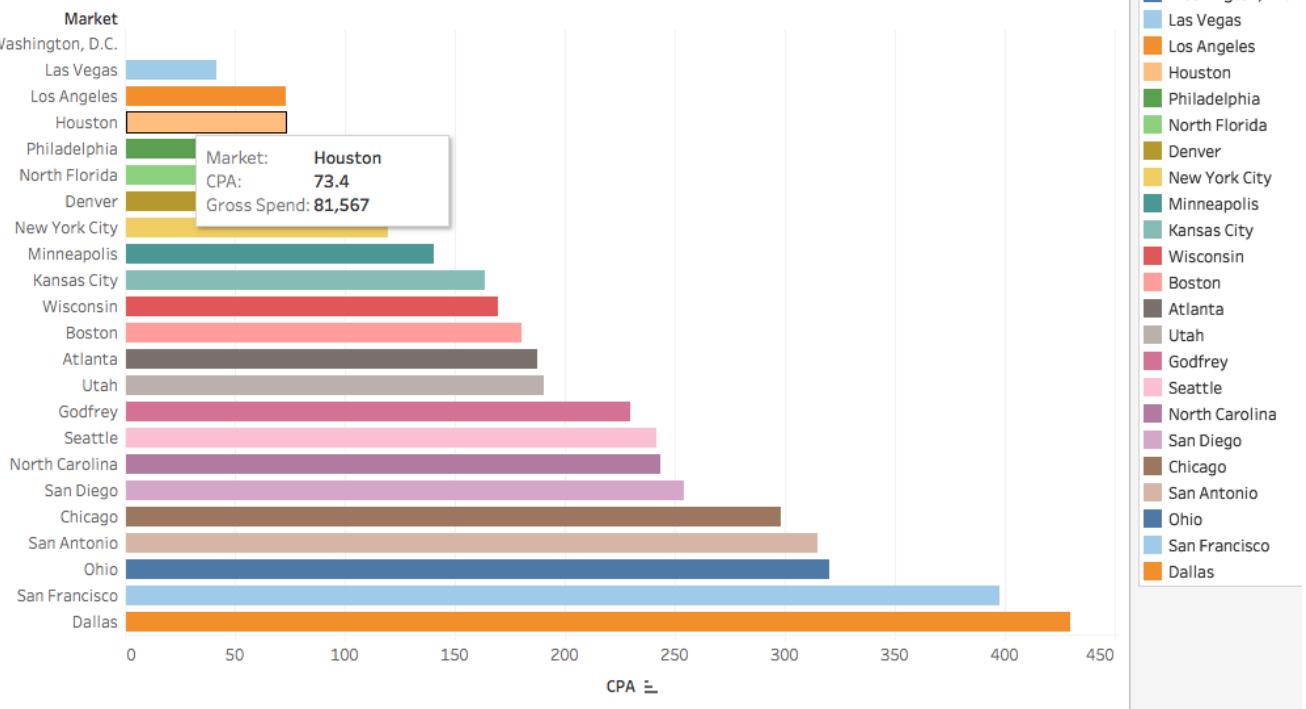
**BI Q8:** What are the greatest opportunities for digital marketing investments? Where have we seen the greatest ROI?

Digital marketing investments are categorized by market as well as by platform.

### Digital marketing trends by market

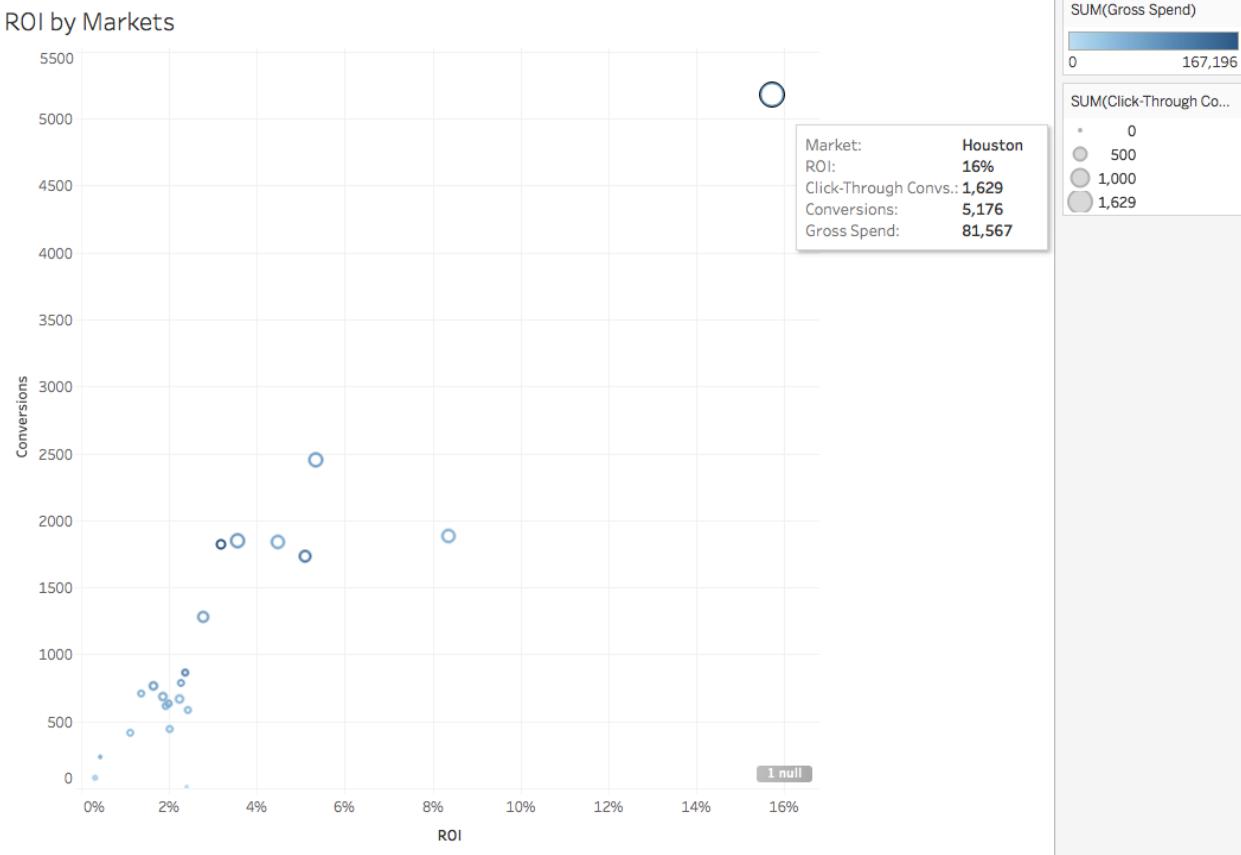
The following three reports shows the digital marketing trends by market. The market with the lowest cost per action (CPA) are the top performing markets with higher conversion rates and greatest return on investments and therefore, are potential opportunities for more investment.

### Top Performing Markets



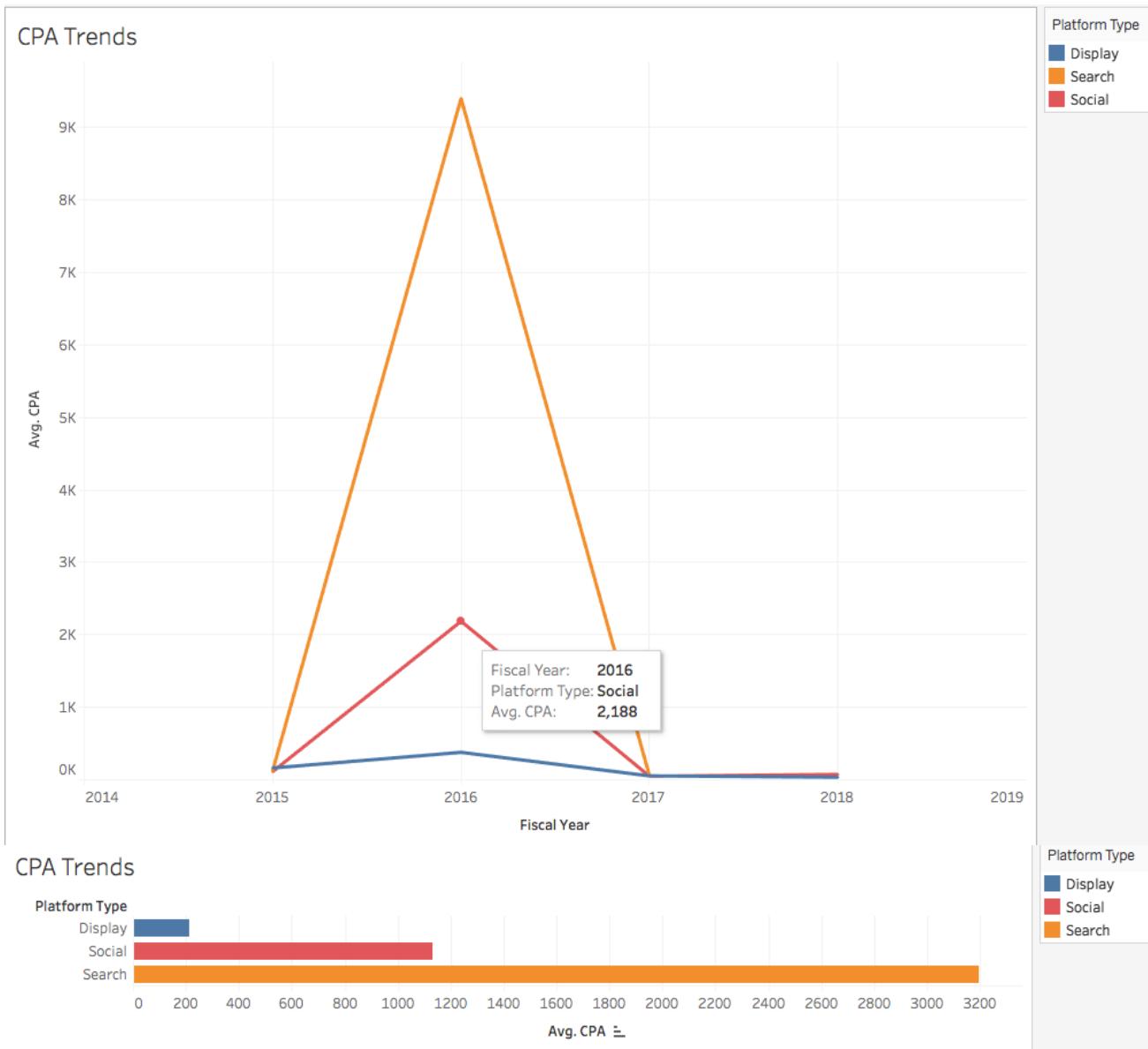
## ROI by Markets

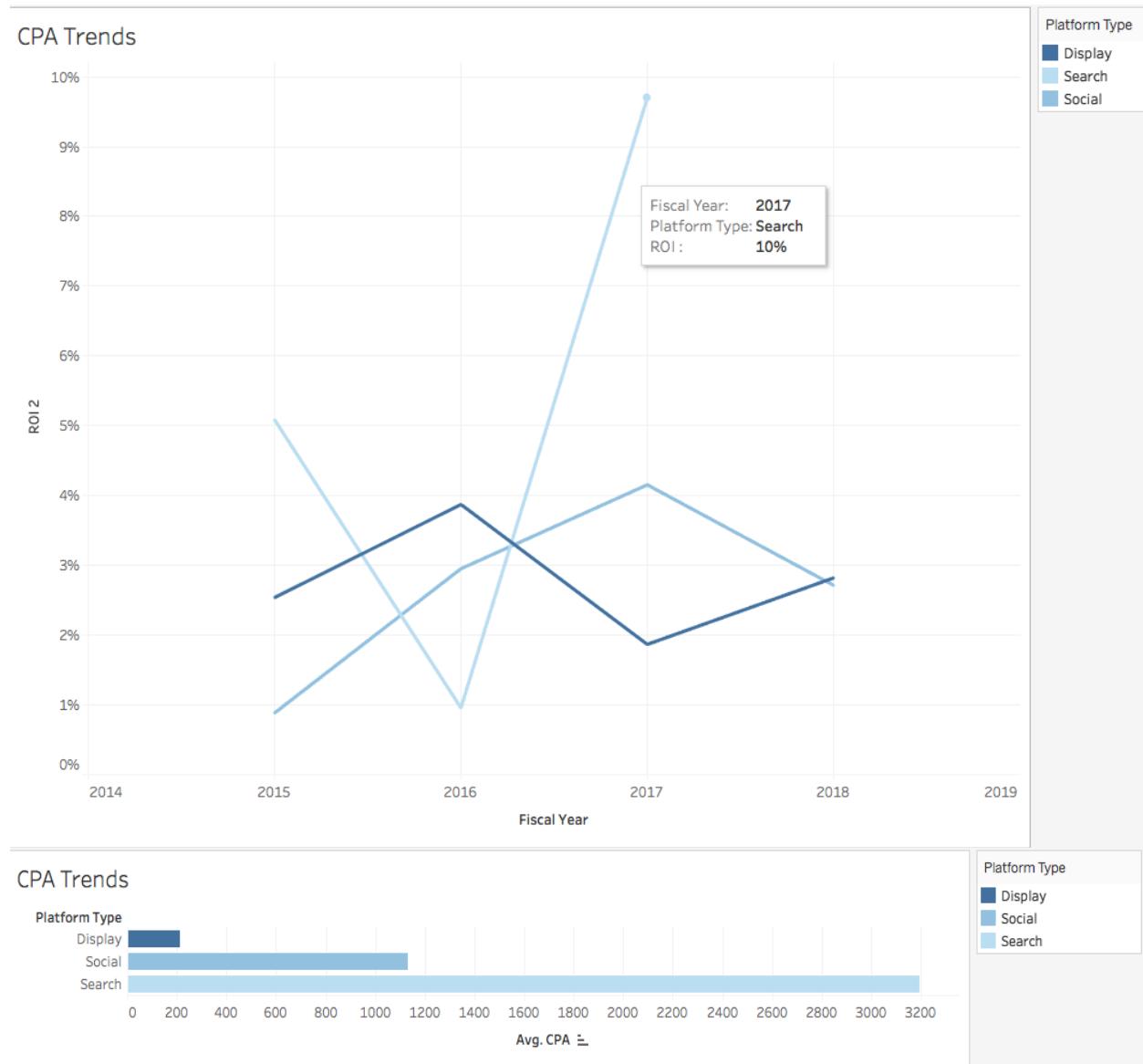
Market	Click-Through		Gross Spend	ROI
	Convs.	F		
Houston	1,629	5,176	81,567	16%
Philadelphia	538	2,447	91,518	5%
New York City	512	1,849	101,538	4%
Minneapolis	461	1,882	68,736	8%
Denver	455	1,838	78,308	4%
Dallas	349	1,733	140,491	5%
Boston	299	1,277	92,372	3%
Chicago	221	1,822	167,196	3%
North Carolina	183	686	72,291	2%
San Diego	177	768	93,090	2%
Utah	158	670	57,203	2%
Godfrey	139	614	55,807	2%
Ohio	131	442	44,330	2%
North Florida	127	414	37,016	1%
Seattle	124	632	69,750	2%
Wisconsin	113	584	47,742	2%
San Francisco	109	863	113,374	2%
Los Angeles	107	707	51,746	1%
Kansas City	101	786	70,082	2%
San Antonio	40	78	24,574	0%
Washington, D.C.	0	0	0	
Atlanta		234	53,807	0%
Las Vegas		10	416	2%



## Digital marketing trends by platform

The following reports shows the digital marketing trends by platform and platform type over the years. The platforms with the lowest cost per action (CPA) are the top performing platforms with higher conversion rates and greatest return on investments and therefore, are potential opportunities for more investment.

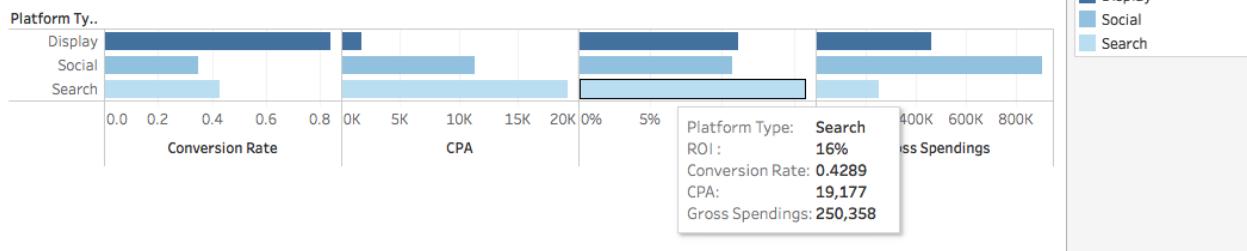




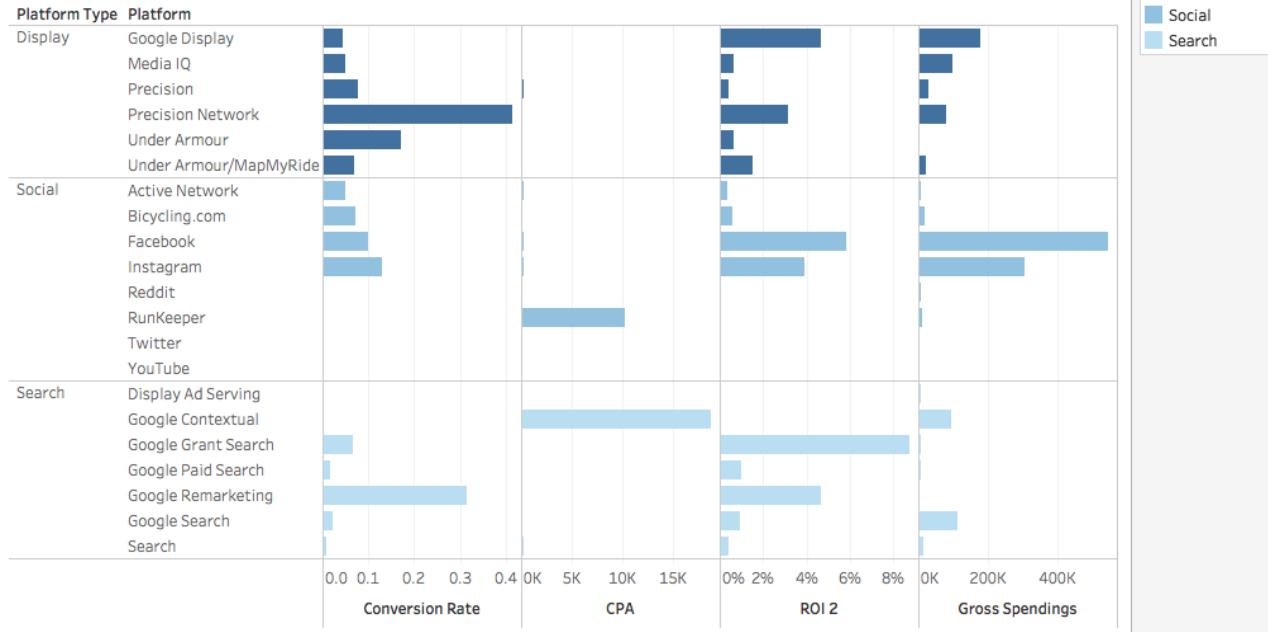
## ROI by platform

Platform Type	CPA	Conversion Rate	Gross Spendings	ROI
Display	1,712	1	460,540	11%
Social	11,298	0	901,291	11%
Search	19,177	0	250,358	16%

## ROI by platform



## ROI by platform type



**BI Q9:** Once someone is registered, what tactics and behaviors drive fundraising, and at what times leading up to the event?

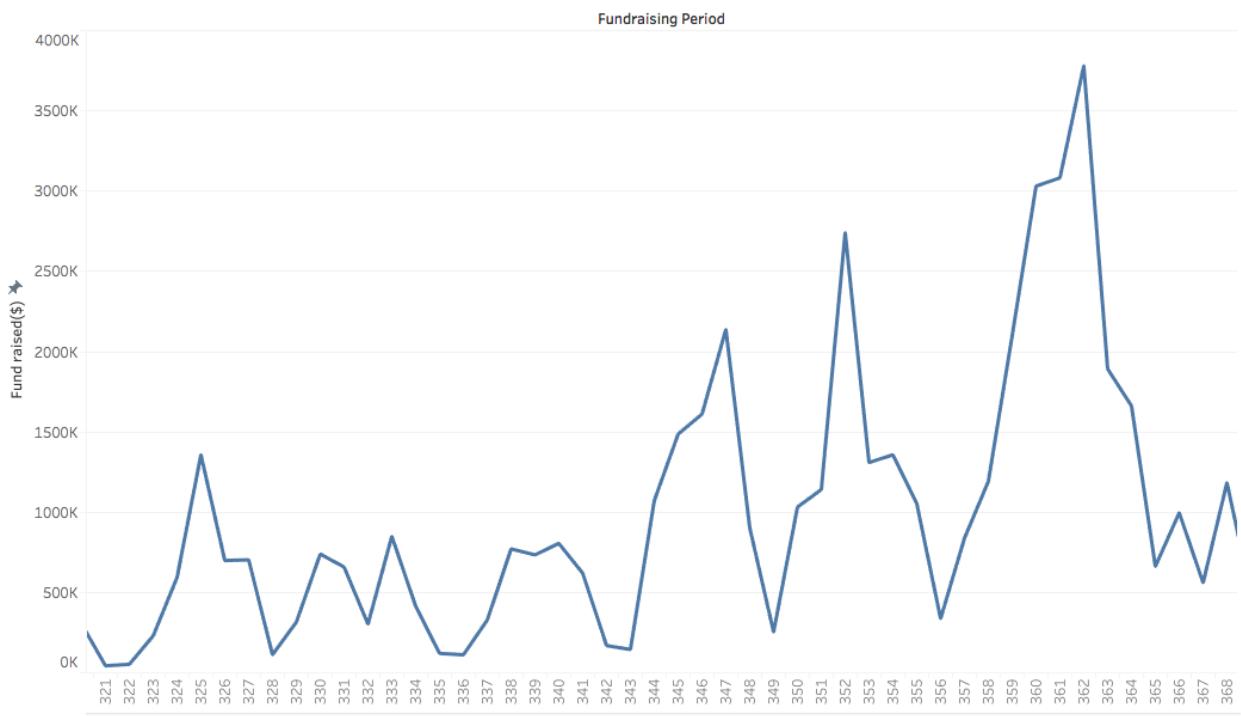
The following two reports shows the fundraising behavior- the donations raised by the participant or a team between the registration date and the event date.

## Fundraising Behavior

Registration Date	Event Date	Fundraising Period	
10/5/2011	10/6/2012	367	67,580
10/10/2011	10/6/2012	362	4,790
10/11/2011	10/6/2012	361	149,393
10/12/2011	10/6/2012	360	100
10/13/2011	10/6/2012	359	7,290
10/14/2011	10/6/2012	358	430
10/16/2011	10/13/2012	363	51,085
10/17/2011	10/6/2012	355	1,400
	10/13/2012	362	82,007
10/18/2011	10/6/2012	354	2,497
	10/13/2012	361	52,629
10/19/2011	10/6/2012	353	9,195
	10/13/2012	360	12,970
10/20/2011	10/6/2012	352	1,125
	10/13/2012	359	32,359
10/21/2011	10/13/2012	358	28,435
10/22/2011	10/13/2012	357	14,896
10/23/2011	10/6/2012	349	1,360
	10/13/2012	356	1,818
10/24/2011	10/6/2012	348	5,275
	10/13/2012	355	613
10/25/2011	10/13/2012	354	290
10/26/2011	10/6/2012	346	0
10/27/2011	10/13/2012	352	300
11/3/2011	10/6/2012	338	3,350
11/4/2011	10/6/2012	337	7,985
11/7/2011	10/6/2012	334	56,103
11/8/2011	10/6/2012	333	11,186

0M      5M  
 Fund Raised

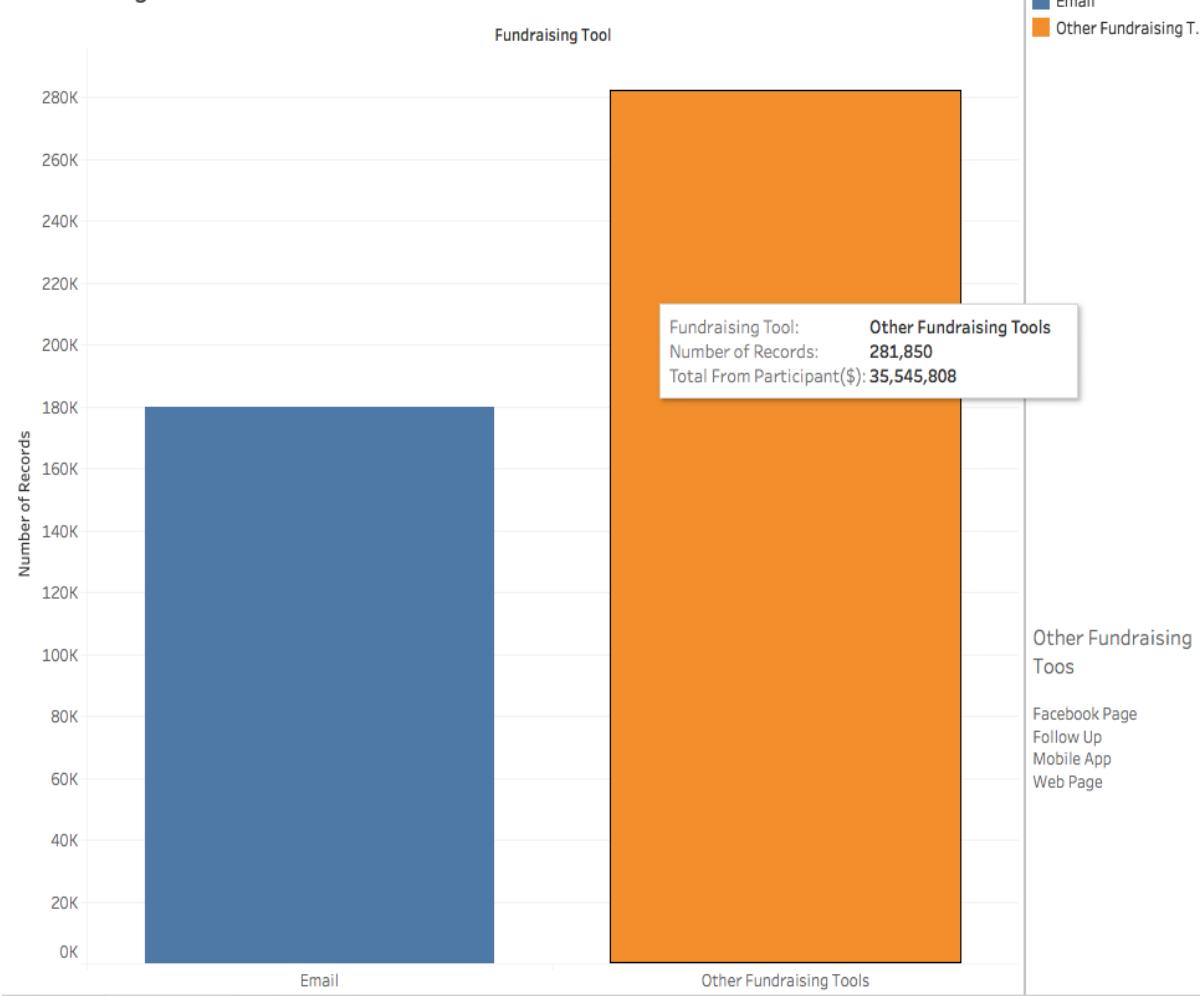
Funds raised after registration  
wrt days before the event



**BI Q10:** What behavioral data do you see about usage of our fundraising tools and how it may or may not relate to performance of top fundraisers?

This report shows the usage of Bike MS' fundraising tools to raise the donations.

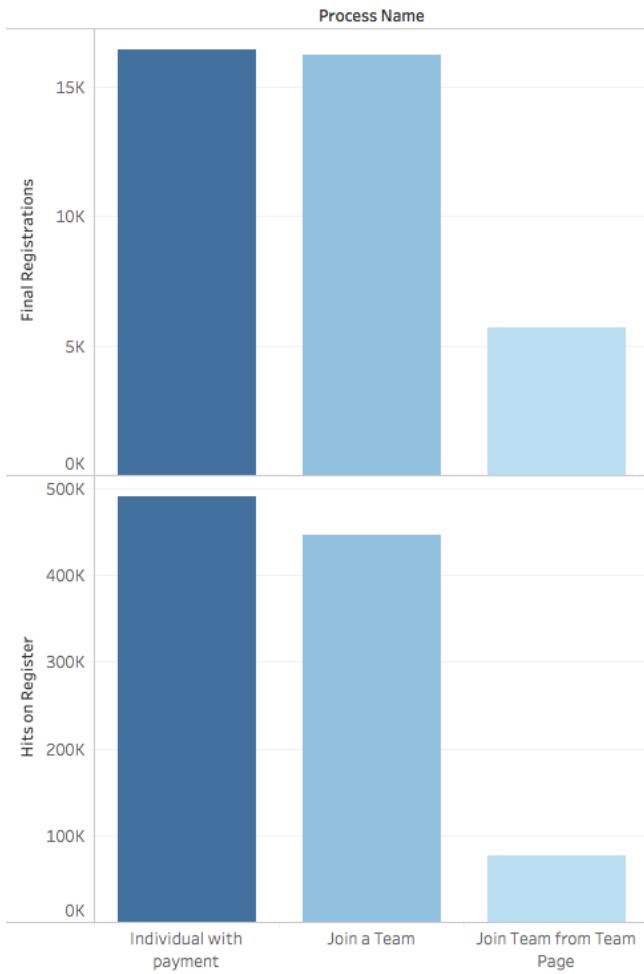
#### Fundraising Tool



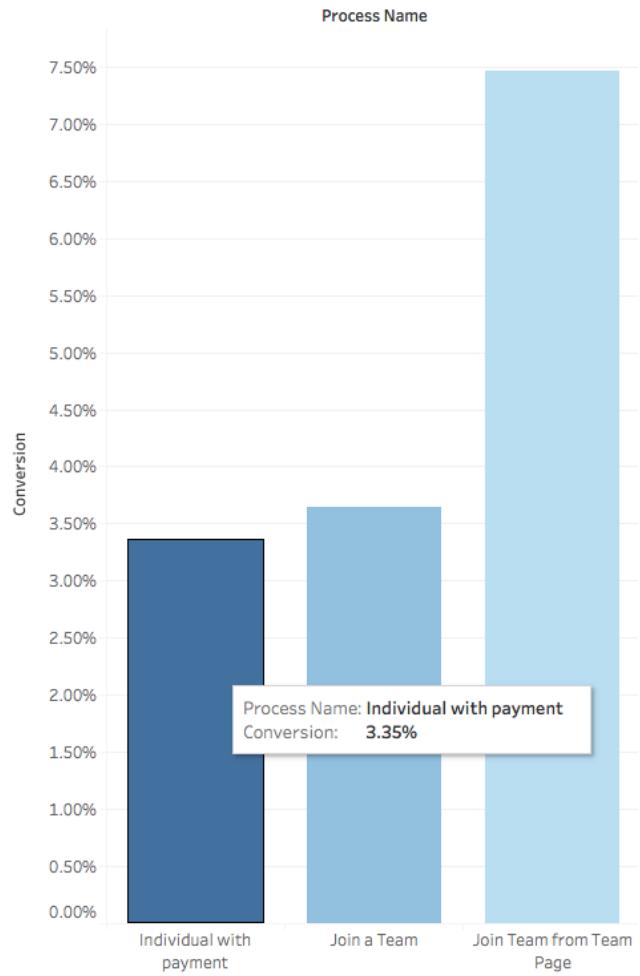
**BI Q11:** Despite increasing our digital advertising spend, acquisition continues to trend downward overall. Why? Is it an issue of needing more traffic, better targeting, or a conversion rate issue that needs to be addressed through the registration process? What can we do to reverse the trend?

This report shows the registration trend: registration hits vs final registration i.e. the conversion rates for three ways in which a participant can register.

Num of Registrations



Conversion Rate



As analyzed in Appendix 1, it's a conversion rate issue that needs to be addressed through the registration process.

Suggestions to reverse the trends:

1. Simplify the registration process- single web-page for registration. Also, there are three ways a participant can register for the event- as an individual, as a team, or join an existing team via search. This process can also be simplified and laid out clearly.
2. Show clear results and feedback to the incoming user
3. Make it more user friendly

## OTHER INSIGHTS

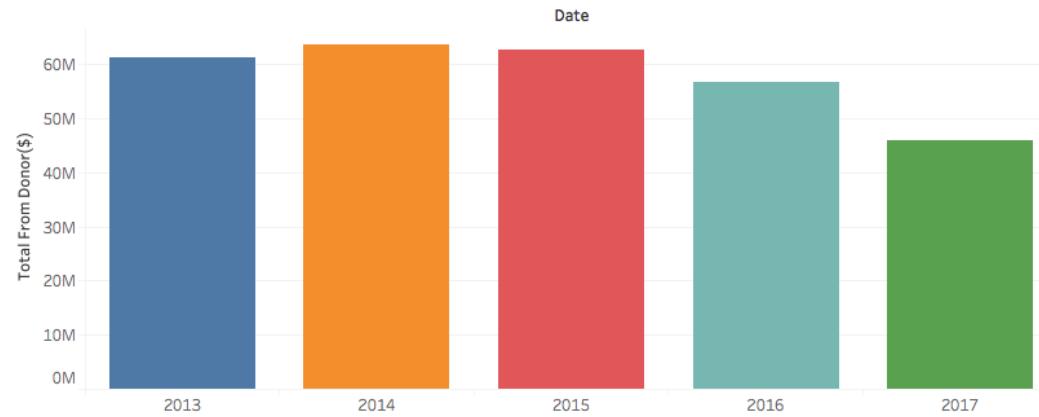
**BI Q12:** As we studied this data, following things came up as an insight into the operations that the questions above do not capture.

### Events

This report shows the revenue raised in each event based on the donations made by the donor and the donations raised by the participants.

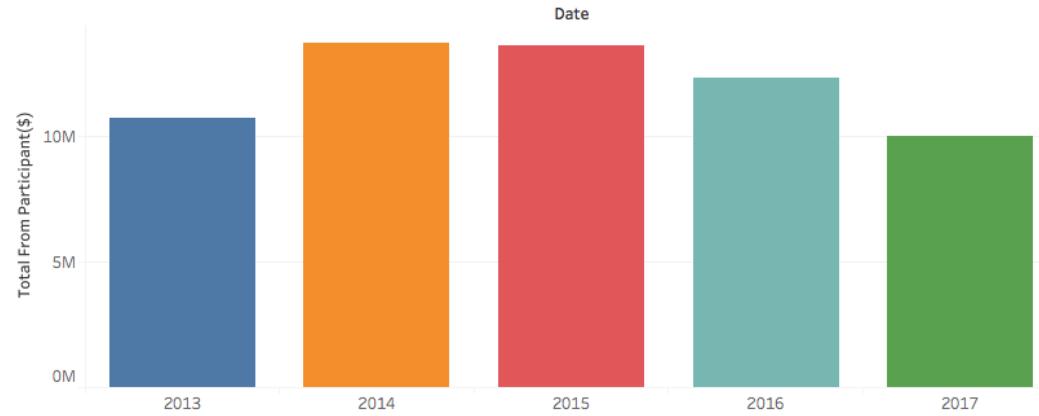
#### Event Revenue

##### From Donors

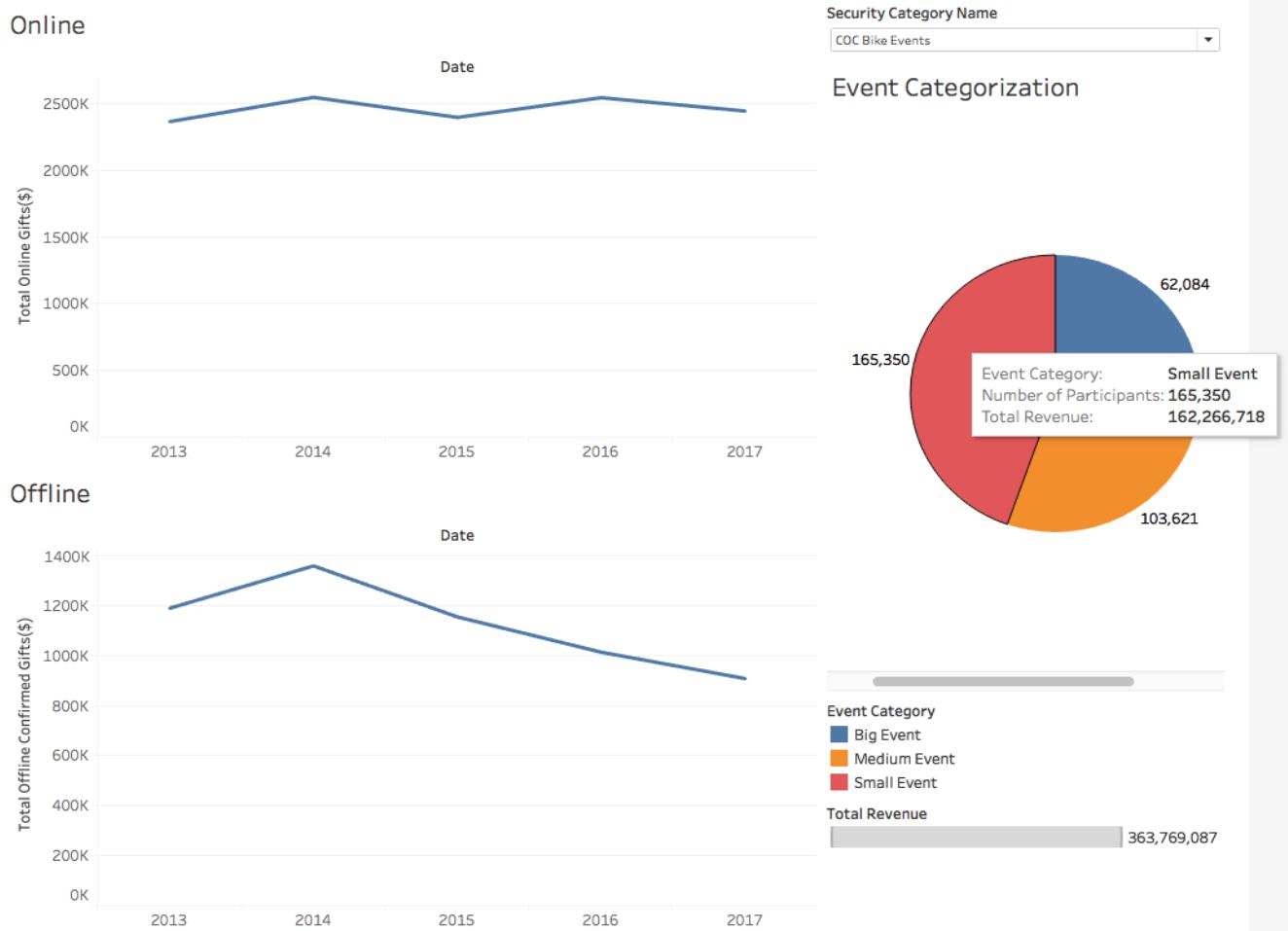


Security Category Name
(All)
Enter search text
(All)
AKA Bike Events
ALC Bike Events
ARR Bike Events
AZA Bike Events
CAL Bike Events
CAN Bike Events
CAS Bike Events
COC Bike Events
CTN Bike Events
DCW Bike Events
DED Bike Events
FLC Bike Events
FLN Bike Events
FLS Bike Events
GAA Bike Events
HIH Bike Events
ILD Bike Events
INI Bike Events
KSG Bike Events
KSS Bike Events
KYW Bike Events
LAM Bike Events
MAM Bike Events
MDM Bike Events
MIG Bike Events
MNM Bike Events
MOS Bike Events
NCC Bike Events
NCT Bike Events

##### From Participants



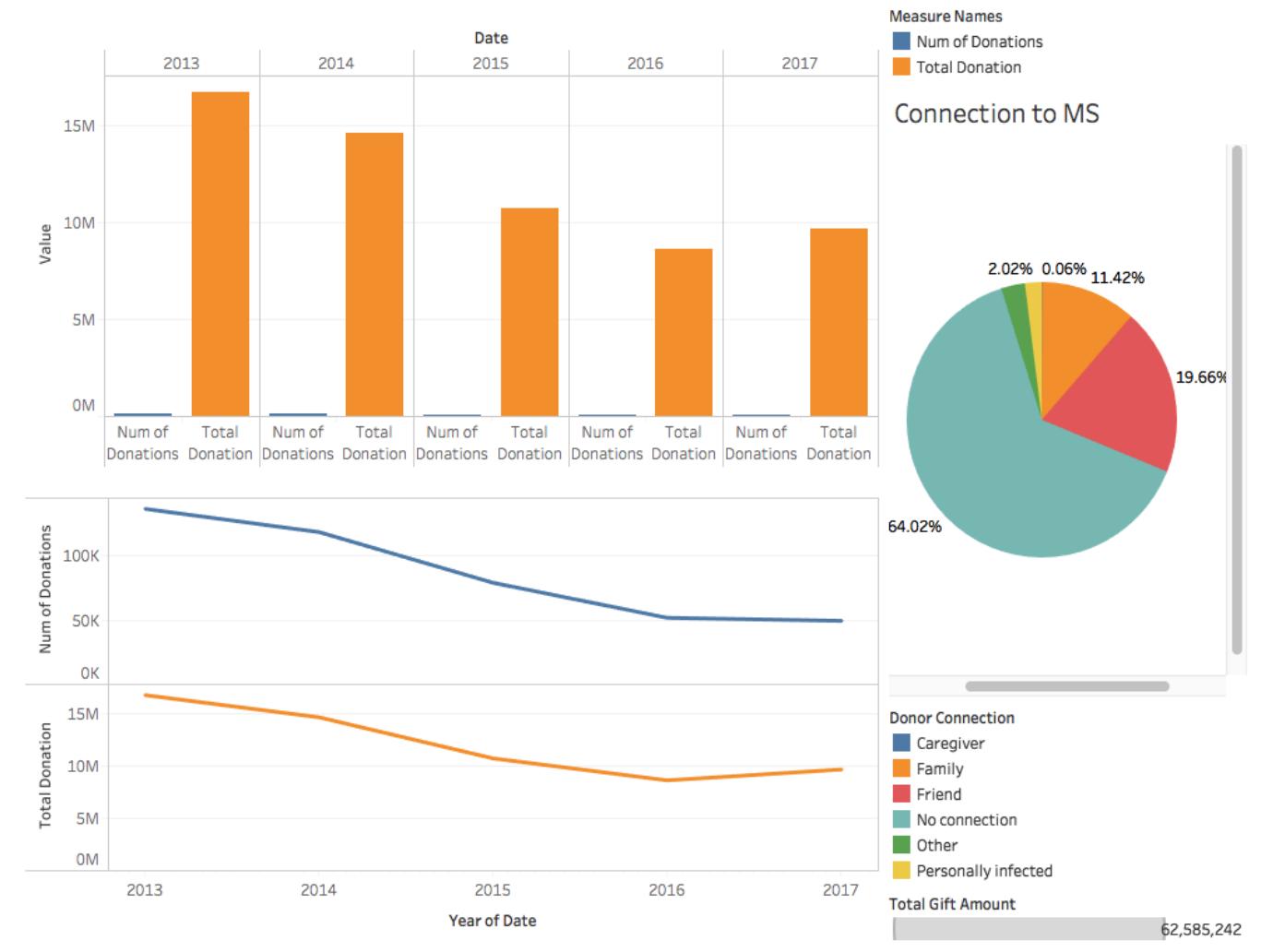
This report shows the distribution of donations by channel type- online vs offline, categorized by events.



## Donations

This report shows the donation made by the donors over the years. As seen, donor connection to MS is not a driving factor for people making donations to the event.

### Donation Trends



## Metadata Documentation

### External data sources Metadata

Metadata	Description
Name	Industry
Brief Description	The industry each participating corporate teams belongs to.
Attributes	<ul style="list-style-type: none"><li>• Company (PK)</li><li>• Industry</li></ul>
Load Frequency	Yearly

Source: *Forbes, S&P*

Metadata	Description
Name	Competitors
Brief Description	The information about the competing events for Bike MS.
Attributes	<ul style="list-style-type: none"><li>• Competitor_Key (PK)</li><li>• Competitor_Name</li><li>• UpperMoneyRange</li><li>• LowerMoneyRange</li><li>• CorporateName</li><li>• Industry</li></ul>
Load Frequency	Yearly

Source:  
*ADA – Tour de Cure*  
*Best Buddies*  
*JDRF*  
*PedalTheCause*

*St. Jude Children's Research Hospital*

## Dimensions Metadata

Dimension Metadata	Description
Name	Dim_Team
Brief Description	The bike team that participates in the Bike MS events to raise fundings.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Team_Key: SCD Type-1</li> <li>● TeamName: SCD Type-1</li> <li>● TeamCreationDate: SCD Type-1</li> <li>● TeamCaptainContactID: SCD Type-1</li> <li>● TeamDivision: SCD Type-1</li> <li>● NationalTeam: SCD Type-1</li> <li>● Company: SCD Type-1</li> <li>● Industry: SCD Type-1</li> <li>● isCompetitor: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_Events
Brief Description	The fundraising event that the bike teams join.
Conformed Dimension	Yes
Hierarchy	None

Change Rules	<ul style="list-style-type: none"> <li>Event_Key: SCD Type-1</li> <li>PublicEventName: SCD Type-1</li> <li>InternalEventName: SCD Type-1</li> <li>SecurityCategoryName: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_Participants
Brief Description	The team members of the bike team.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>Participant_Key: SCD Type-1</li> <li>ParticipantGender: SCD Type-1</li> <li>Participant_Connection_to_MS: SCD Type-1</li> <li>Participant_Registration_Type: SCD Type-2</li> <li>RegistrationActiveStatus: SCD Type-1</li> <li>IsTeamCaptain: SCD Type-1</li> <li>ParticipantEmployer: SCD Type-1</li> <li>ParticipantOccupation: SCD Type-1</li> <li>ParticipantContactID: SCD Type-1</li> <li>Effective_Date: SCD Type-1</li> <li>Current_Flag: SCD Type-1</li> <li>ParticipantType: SCD Type-2</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_Donors

Brief Description	The people that donate money or gift items to the fundraising event.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Donor_Key: SCD Type-1</li> <li>● DonorMemberID: SCD Type-1</li> <li>● DonorGender: SCD Type-1</li> <li>● DonorOptoutDate: SCD Type-1</li> <li>● DonorOptoutMethod: SCD Type-1</li> <li>● DonorConnectionToMS: SCD Type-1</li> <li>● DonorEmployer: SCD Type-1</li> <li>● DonorConsID: SCD Type-1</li> <li>● DonorEmailStatus: SCD Type-1</li> <li>● DonorAcceptEmail: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_MarketAdvertisements
Brief Description	The advertisements that NMSS put for the events in the market.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Market_Key: SCD Type-1</li> <li>● MarketName: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_PlatformAdvertisements
Brief Description	The advertisements that NMSS put for the events on the platform.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Platform_Key: SCD Type-1</li> <li>● PlatformName: SCD Type-1</li> <li>● PlatformType: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_Affiliates
Brief Description	Original 3-byte chapter code.
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Affiliate_Key: SCD Type-1</li> <li>● Chapter_Name: SCD Type-1</li> <li>● Affiliate_Code_mrg: SCD Type-1</li> <li>● Affiliate_Code</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_GoogleAnalytics
Brief Description	The process of Google Analytics
Conformed Dimension	Yes
Hierarchy	None
Change Rules	<ul style="list-style-type: none"> <li>● Process_Key: SCD Type-1</li> <li>● Process_Name: SCD Type-1</li> </ul>
Load Frequency	Daily

Dimension Metadata	Description
Name	Dim_Date
Brief Description	The calendar date.
Conformed Dimension	Yes
Hierarchy	Day < Week < Month < Season < Year
Change Rules	<ul style="list-style-type: none"> <li>● Date_Key: SCD Type-1</li> <li>● Date: SCD Type-1</li> <li>● Month: SCD Type-1</li> <li>● Year: SCD Type-1</li> <li>● Day: SCD Type-1</li> <li>● WeekNo: SCD Type-1</li> <li>● MonthName: SCD Type-1</li> <li>● Season: SCD Type-1</li> <li>● IsWeekend: SCD Type-1</li> </ul>
Load Frequency	Daily

## Facts Metadata

Fact Table Metadata	Description
Name	Fact_Participation
Type	Accumulating Periodic Snapshot Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>● Events</li> <li>● Team</li> <li>● Participants</li> <li>● Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>● NumberofEmailsSent: sum of the emails sent.</li> <li>● TotalFromParticipants: sum of donations raised from the participant.</li> <li>● ParticipantGoal: fundraising goal of the participant.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>● NumberofEmailsSent: additive</li> <li>● TotalFromParticipants: additive</li> <li>● ParticipantGoal: additive</li> </ul>
Format of the Measures/Facts	<ul style="list-style-type: none"> <li>● NumberofEmailsSent: INT[10]</li> <li>● TotalFromParticipants: INT[10]</li> <li>● ParticipantGoal: INT[10]</li> </ul>
Load Frequency	Daily

Fact Table Metadata	Description
Name	Fact_Donations
Type	Transaction Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>● Donor</li> <li>● Affiliate</li> </ul>

	<ul style="list-style-type: none"> <li>Events</li> <li>Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>TotalGiftAmount: sum of donations from the donor.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>TotalGiftAmount: additive</li> </ul>
Format of the Measures/Facts	<ul style="list-style-type: none"> <li>TotalGiftAmount: INT[10]</li> </ul>
Load Frequency	Daily

Fact Table Metadata	Description
Name	Fact_Events
Type	Transaction Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>Events</li> <li>Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>Event_Goal: the fundraising goal of the event.</li> <li>AverageTeamSize: the average size of the teams in an event.</li> <li>NumberOfTeams: the total amount of the teams that participate in the event.</li> <li>NumberOfParticipants: the total amount of the participants that participate in the event.</li> <li>EventsGoalsPerPerson: the average number of the fundraising goal of each person.</li> <li>ActiveRegistration: the total amount of the participants</li> </ul>

	<p>that register actively.</p> <ul style="list-style-type: none"> <li>● InactiveRegistration: the total amount of the participants that register inactively.</li> <li>● TotalOfflineGifts: the total amount of the offline gifts.</li> <li>● TotalOnlineGifts: the total amount of the online gifts.</li> <li>● TotalRevenue: the total amount of the revenue of the event.</li> <li>● TotalFromParticipants: the total amount of funds raised from the participants.</li> <li>● TotalFromDonors: the total amount of funds raised from the donors.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>● Event_Goal: additive</li> <li>● AverageTeamSize: non-additive</li> <li>● NumberOfTeams: additive</li> <li>● NumberOfParticipants: additive</li> <li>● EventsGoalsPerPerson: non-additive</li> <li>● ActiveRegistration: additive</li> <li>● InactiveRegistration: additive</li> <li>● TotalOfflineGifts: additive</li> <li>● TotalOnlineGifts: additive</li> <li>● TotalRevenue: additive</li> <li>● TotalFromParticipants: additive</li> <li>● TotalFromDonors: additive</li> </ul>
Format of the Measures/Facts	<ul style="list-style-type: none"> <li>● Event_Goal: INT[10]</li> <li>● AverageTeamSize: INT[3]</li> <li>● NumberOfTeams: INT[5]</li> <li>● NumberOfParticipants: INT[6]</li> <li>● EventsGoalsPerPerson: INT[7]</li> <li>● ActiveRegistration: INT[5]</li> <li>● InactiveRegistration: INT[5]</li> <li>● TotalOfflineGifts: INT[10]</li> <li>● TotalOnlineGifts: INT[10]</li> <li>● TotalRevenue: INT[10]</li> <li>● TotalFromParticipants: INT[10]</li> <li>● TotalFromDonors: INT[10]</li> </ul>

Load Frequency	Daily
----------------	-------

Fact Table Metadata	Description
Name	Fact_MarketAdvertisement_Analysis
Type	Transaction Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>MarketAdvertisements</li> <li>Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>DeliveredImpressions: the total number of the delivered impressions of the advertisement.</li> <li>DeliveredClicks: the total number of the delivered clicks of the advertisement.</li> <li>CTR: the ratio of users who click on a specific link to the number of total users who view the advertisement.</li> <li>Click-ThroughConversions: the total number of the clicks through conversions.</li> <li>View-ThroughConversions: the total number of the views through conversions.</li> <li>TotalConversions: click-through conversions plus view through conversions.</li> <li>ConversionRate: the percentage of visitors to the advertisement that complete a desired goal (a conversion) out of the total number of visitors.</li> <li>GrossSpend: the gross spend on putting the advertisement.</li> <li>CPA: (cost per acquisition) How much the business pays in order to attain a conversion.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>DeliveredImpressions: additive</li> <li>DeliveredClicks: additive</li> <li>CTR: additive</li> <li>Click-ThroughConversions: additive</li> </ul>

	<ul style="list-style-type: none"> <li>• View-ThroughConversions: additive</li> <li>• TotalConversions: additive</li> <li>• ConversionRate: non-additive</li> <li>• GrossSpend: additive</li> <li>• CPA: non-additive</li> </ul>
Format of the Measures/Facts	<ul style="list-style-type: none"> <li>• DeliveredImpressions: INT[7]</li> <li>• DeliveredClicks: INT[7]</li> <li>• CTR: INT[10]</li> <li>• Click-ThroughConversions: INT[7]</li> <li>• View-ThroughConversions: INT[7]</li> <li>• TotalConversions: INT[10]</li> <li>• ConversionRate: FLOAT</li> <li>• GrossSpend: INT[7]</li> <li>• CPA: FLOAT</li> </ul>
Load Frequency	Daily

Fact Table Metadata	Description
Name	Fact_PlatformAdvertisement_Analysis
Type	Transaction Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>• PlatformAdvertisements</li> <li>• Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>• DeliveredImpressions: the total number of the delivered impressions of the advertisement.</li> <li>• DeliveredClicks: the total number of the delivered clicks of the advertisement.</li> <li>• CTR: the ratio of users who click on a specific link to the number of total users who view the advertisement.</li> <li>• Click-ThroughConversions: the total number of the clicks through conversions.</li> </ul>

	<ul style="list-style-type: none"> <li>• View-ThroughConversions: the total number of the views through conversions.</li> <li>• TotalConversions: click-through conversions plus view through conversions.</li> <li>• ConversionRate: the percentage of visitors to the advertisement that complete a desired goal (a conversion) out of the total number of visitors.</li> <li>• GrossSpend: the gross spend on putting the advertisement.</li> <li>• CPA: (cost per acquisition) How much the business pays in order to attain a conversion.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>• DeliveredImpressions: additive</li> <li>• DeliveredClicks: additive</li> <li>• CTR: additive</li> <li>• Click-ThroughConversions: additive</li> <li>• View-ThroughConversions: additive</li> <li>• TotalConversions: additive</li> <li>• ConversionRate: non-additive</li> <li>• GrossSpend: additive</li> <li>• CPA: non-additive</li> </ul>
Format of the Measures/Facts	<ul style="list-style-type: none"> <li>• DeliveredImpressions: INT[7]</li> <li>• DeliveredClicks: INT[7]</li> <li>• CTR: INT[10]</li> <li>• Click-ThroughConversions: INT[7]</li> <li>• View-ThroughConversions: INT[7]</li> <li>• TotalConversions: INT[10]</li> <li>• ConversionRate: FLOAT</li> <li>• GrossSpend: INT[7]</li> <li>• CPA: FLOAT</li> </ul>
Load Frequency	Daily

Fact Table Metadata	Description
Name	Fact_GoogleAnalytics

Type	Transaction Fact Table
Relevant Dimensions	<ul style="list-style-type: none"> <li>● GoogleAnalytics</li> <li>● Date</li> </ul>
Degenerate Dimension	None
Measures/Facts Description	<ul style="list-style-type: none"> <li>● HitsOnTeamPage: the total number of the hits on team page.</li> <li>● HitsOnRegister: the total number of the hits on registration page.</li> <li>● HitsonSearchForTeam: the total number of the hits on searching for team page.</li> <li>● HitsOnSearchResults: the total number of the hits on searching results page.</li> <li>● HitsOnSelectOptions: the total number of the hits on selecting options page.</li> <li>● HitsOnProvideDetails: the total number of the hits on providing details page.</li> <li>● HitsOnAgreeToTerms: the total number of the hits on “agree to terms” page.</li> <li>● HitsOnReview: the total number of the hits on review page.</li> <li>● HitsOnPayment: the total number of the hits on payment page.</li> <li>● FinalRegistrations the total number of the hits on final registration page.</li> </ul>
Measures/Facts Type	<ul style="list-style-type: none"> <li>● HitsOnTeamPage: additive</li> <li>● HitsOnRegister: additive</li> <li>● HitsonSearchForTeam: additive</li> <li>● HitsOnSearchResults: additive</li> <li>● HitsOnSelectOptions: additive</li> <li>● HitsOnProvideDetails: additive</li> <li>● HitsOnAgreeToTerms: additive</li> <li>● HitsOnReview: additive</li> <li>● HitsOnPayment: additive</li> <li>● FinalRegistrations: additive</li> </ul>

Format of the Measures/Facts	<ul style="list-style-type: none"><li>• HitsOnTeamPage: INT[7]</li><li>• HitsOnRegister: INT[7]</li><li>• HitsonSearchForTeam: INT[7]</li><li>• HitsOnSearchResults: INT[7]</li><li>• HitsOnSelectOptions: INT[7]</li><li>• HitsOnProvideDetails: INT[7]</li><li>• HitsOnAgreeToTerms: INT[7]</li><li>• HitsOnReview: INT[7]</li><li>• HitsOnPayment: INT[7]</li><li>• FinalRegistrations: INT[7]</li></ul>
Load Frequency	Daily

## Appendices

### Appendix 1

To reason out the downward trending digital advertising spend vs acquisition, we analyzed the registration process of Best Buddies and compared it with Bike MS.

1. The registration process for Bike MS is a stepwise, tedious and time-consuming process spanning multiple pages as compared to Best Buddies single sign-in webpage.

# Register

If you registered for a past Best Buddies event, [click here](#) to use your existing username and password to register. Once you have logged in, click 'Agree' on the waiver to register.

## Personal Information

First Name \*

Last Name \*

Email \*

Address Line 1 \*

City \*

Province/State \*

Postal Code \*

Phone \*

Date of Birth \*

Gender \*

## Registration

Houston Walker Registration

Donate or raise just \$50 and you'll get an exclusive Best Buddies Friendship Walk T-shirt.

Emergency Contact Name

Emergency Contact Phone

T-Shirt Size

Are you joining a school team?

**Would you like to participate on a team? Skip this step if you'd prefer to register as an individual:**

**Create a team**      **Join a team**

Team name:  Team name:

**CREATE**      Show All **SEARCH**

**Would you like to start your fundraising by making a self-donation? (not required)**

**Donation**

\$500.00       \$100.00       \$75.00       \$50.00       \$25.00

Donation Amount:

Maximize your gift to Best Buddies by covering our 2.9% credit card processing fee. - \$0.00

**Honor Roll**

Show on honor roll

Name

Message

Show Amount

**User Account**

Username:

Password:

Password Confirm:

**SUBMIT**

2. Bike MS mandates the participant to enter a minimum fundraising goal of \$250 before registering, which is not necessarily needed to be achieved in an event. It might be a

potential reason for discouraging participants from signing up. This is optional while registering for Best buddies.

## Bike MS: Ride the Vineyard 2018

- 1 Fundraising Goal
- 2 Participant Details
- 3 Agree to Terms
- 4 Review Registration
- 5 Make Payment

! The participation type you selected has a minimum fundraising goal of \$250.00.

### Your Fundraising Goal

Enter discount code:

Our fundraising minimum is \$250, but you can push it further.

Your Personal Fundraising Goal:

Demonstrate your personal commitment to helping those who live with MS by making the first donation. You'll see how much easier it is to raise money when you give first.

\$35.00

\$50.00

\$100.00

\$200.00

\$250.00

\$500.00

\$1,000.00

Other amount that is meaningful to you:

No additional gift

Yes, make this an anonymous gift.

Yes, you can display the amount of my donation publicly.

Yes, associate my participation with a company.

**Next Step**

Based on our following analysis, it is an issue of conversion rate.

## References

1. Date dimension <https://www.mssqltips.com/sqlservertip/4054/creating-a-date-dimension-or-calendar-table-in-sql-server/>

## Glossary

Term	Definition
<b>BI</b> <b>Business Intelligence</b>	A suite of software tools used primarily by business administrative staff to navigate through the data of the data warehouse. BI tools provide functionality including managed reporting, querying, data analysis, data visualization, etc.
<b>Cube</b>	A multi-dimensional representation of business data in which the cells of the cube contain data measures (i.e. facts) and the edges of the cube represent the data dimensions. Although a cube implies only 3 dimensions in geometry, a data cube may represent any number of dimensions.
<b>Data mart</b>	A subset of the organization's data, focused on a specific subject area or business area.
<b>Data warehouse</b> <b>DW</b>	(1) A collection of data pulled together primarily from operational business systems, structured and tuned for easy access and use by consumers and analysts, especially in support of forecasting and decision-making. (2) A subject-oriented, integrated, time-varying, non-volatile collection of data in support of the management's decision-making process. (John Inmon)
<b>Dimension</b>	A set of attributes, usually hierarchical, that is used to describe an organization's business by constraining and grouping facts. Example dimensions include time, students, faculty, organization, funds, etc.
<b>ETL</b> <b>Extraction, Transformation, Loading</b>	A set of back-end data staging steps that are used to (1) obtain data from operational sources (i.e. the extraction step), (2) cleanse and prepare data for import into the data warehouse (i.e. the transformation step), and (3) actually importing the transformed data into the data warehouse (i.e. the loading step).
<b>Fact</b>	A numeric (or other type of) data element by which an organization measures aspects of its business. The most useful facts are indeed numeric and often additive.  Example facts include dollar amounts (e.g. budget, expenditure, encumbrance, revenue), counts (e.g. headcount, credit hours), etc.
<b>Grain</b>	The meaning of a single record in a fact table.
<b>Metadata</b>	Data about data; any data maintained to support the operation or use of a data warehouse, including business names and definitions of facts, dimensions, attributes, etc.
<b>ODBC</b> <b>Open Database Connectivity</b>	Often used in a Microsoft Windows environment, ODBC is a standard protocol for accessing a majority of database systems, including Oracle, Access, etc.
<b>OLAP</b> <b>On-Line Analytical Processing</b>	A category of database software systems that primarily involves aggregating large amounts of data from a data warehouse environment.
<b>OLTP</b> <b>On-Line Transactional Processing</b>	A category of database software systems that typically involves processing transactions in real time.
<b>RDBMS</b> <b>Relational Database Management System</b>	A database (e.g. Oracle) in which information is represented via tables and relationships between such tables. The term RDBMS is also often used to refer to software that helps to administer a database.
<b>Accumulating snapshot</b>	Describes a fact table that represents the state of affairs at the end of each time period.

<b>Star schema</b>	A collection of dimensions joined together with a single fact table that is used to construct queries against a data warehouse.
<b>SSIS package (SQL Server Integration Services)</b>	SSIS is a platform for data integration and workflow applications. It features a data warehousing tool used for data extraction, transformation, and loading (ETL). The tool may also be used to automate maintenance of SQL Server databases and updates to multidimensional cube data. SSIS packages are a collection of data flows.

<http://www.rpi.edu/datawarehouse/dw-glossary.html>

## Peer Evaluation

1. Divyesh Batra
2. Harikrishnan Subramanyan
3. Nupur Sinha
4. Tanushree Kothari
5. Tsai Mei Chen

Criteria	Maximum Points	1	2	3	4	5
Shows up to meetings	10					
Demonstrates respect for other team members	10					
Level of contribution to the team	20					
Accepts and shares team responsibilities	10					
Actively contributes to project completion	15					
Willing to help other team members	10					
Has a positive attitude towards team	10					
Actively contributes to the preparation of the	15					

report						
Total individual scores	100					