

# Introduction to Data Visualization

## DATA VISUALIZATION



Mr.Fahad Masood Reda

 <https://themis427.wordpress.com>



themisblog



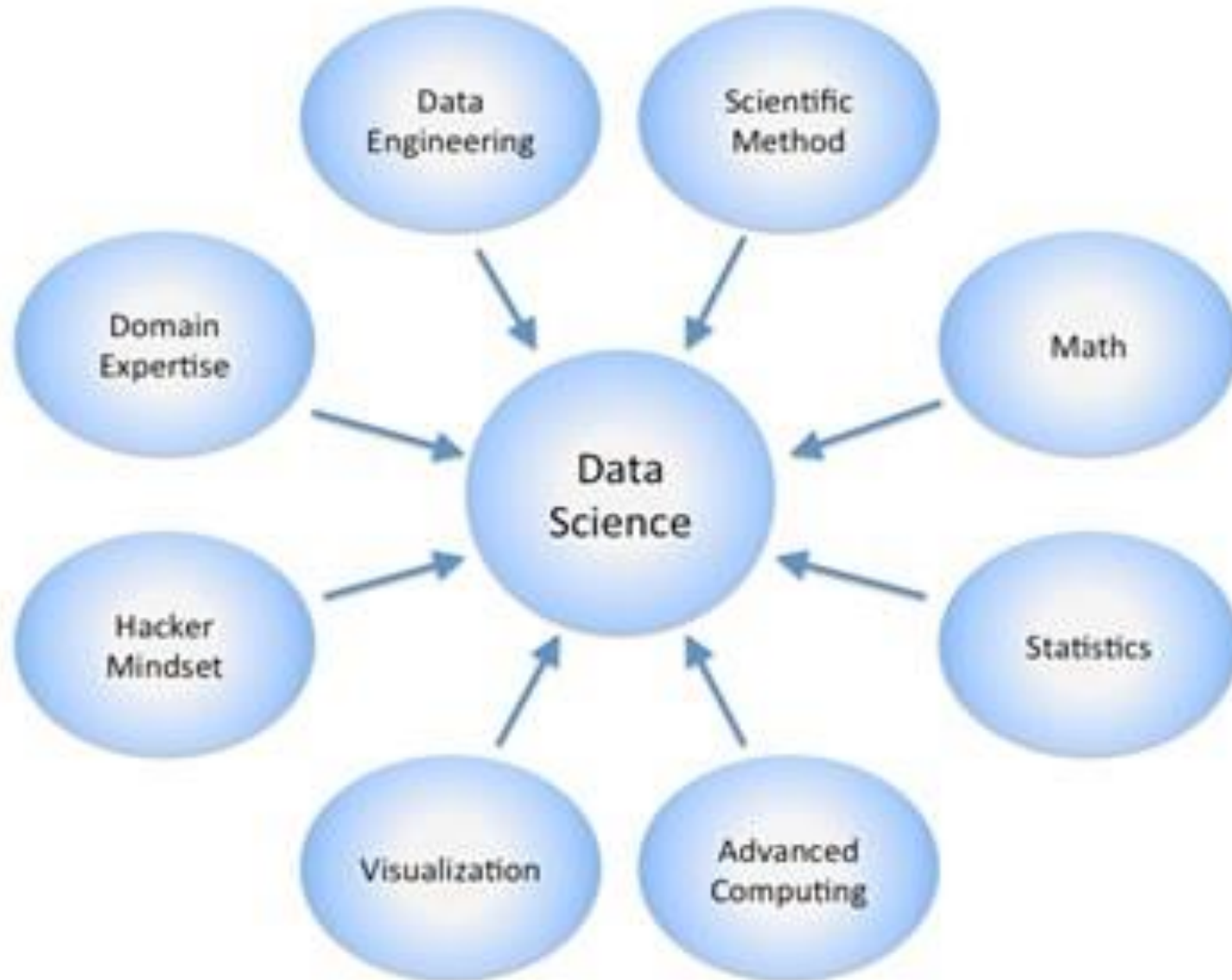
themis427



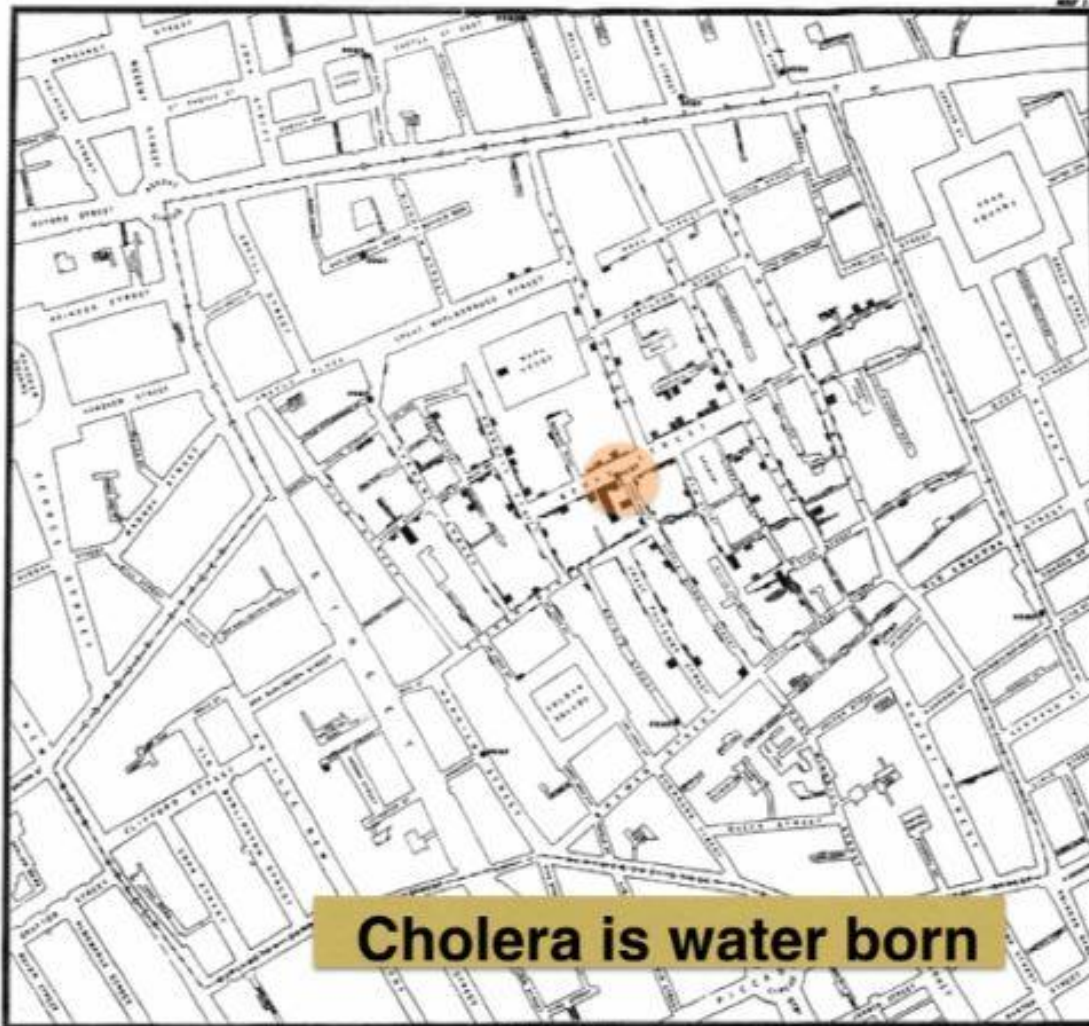
# Learning Objectives

- ▶ Importance of Data Visualization
- ▶ Understanding Data Pyramid (DIKW pyramid)
- ▶ What is Data Visualization
- ▶ Data Visualization Process Explained
- ▶ Why Data Visualization ( Benefits of Data Visualization)
- ▶ Types of data visualization
- ▶ Some Popular Software's for Data Visualization
- ▶ Tableau Software

# Importance of Data Visualization



# Importance of Data Visualization



John Snow

# Importance of Data Visualization

“Data is just like crude.

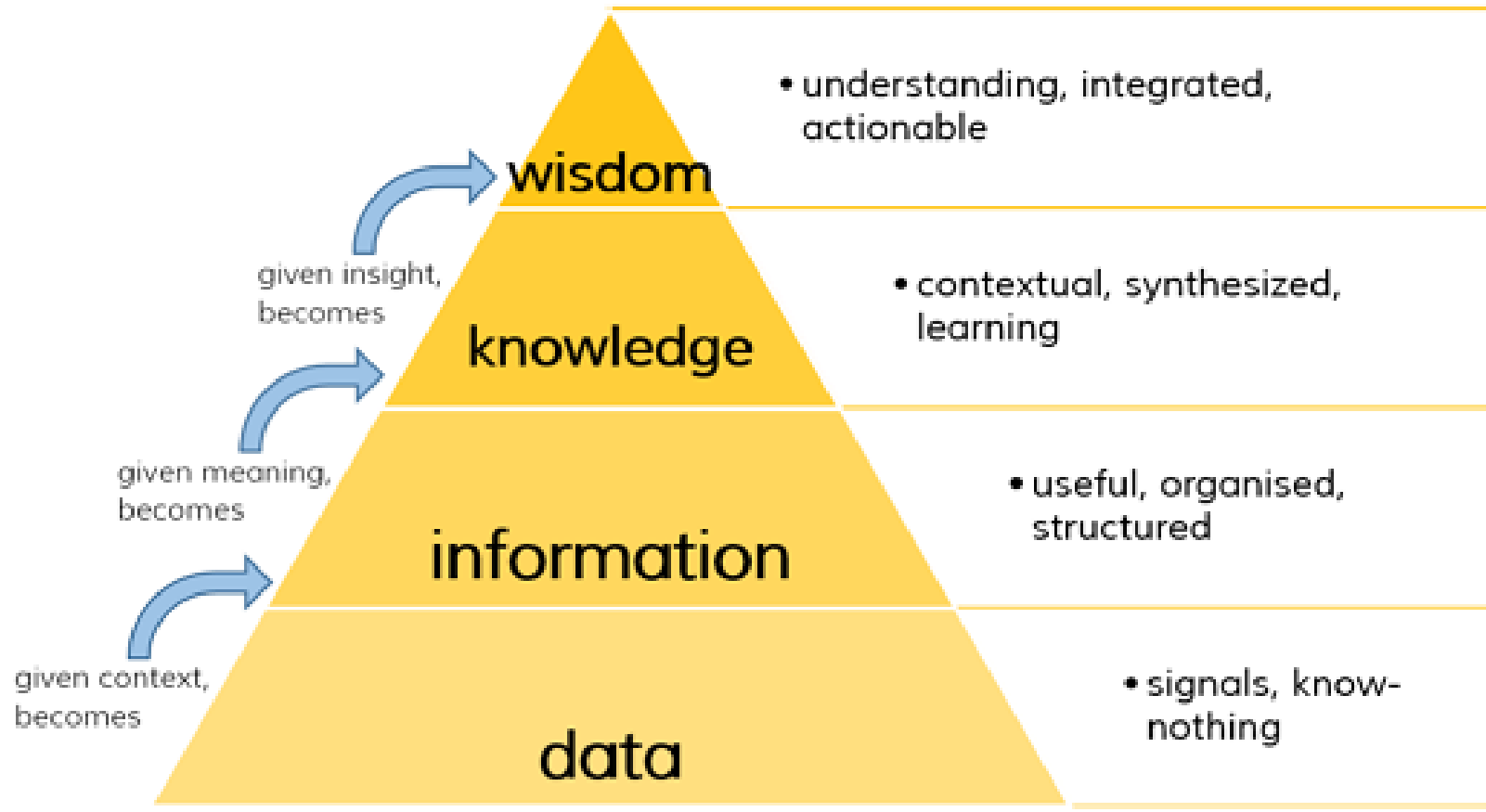
It’s valuable, but if unrefined it cannot really be used.

It has to be changed into gas, plastic, chemicals, etc., to create a valuable entity that drives profitable activity;

so must data be broken down and analysed for it to have value.”

-Michael Palmer

# Data, Information, Knowledge, wisdom



# Data, Information, Knowledge , wisdom

Data is unprocessed facts and figures without any added interpretation or analysis. "The price of crude oil is \$80 per barrel."

Information is data that has been interpreted so that it has meaning for the user. "The price of crude oil has risen from \$70 to \$80 per barrel" gives meaning to the data and so is said to be information to someone who tracks oil prices.

Knowledge is a combination of information, experience and insight that may benefit the individual or the organization. "When crude oil prices go up by \$10 per barrel, it's likely that petrol prices will rise by 2% per liter" is knowledge.

# Data, Information, Knowledge , wisdom

البيانات: عبارة عن كلمات مفصولة عن بعضها ، كل كلمة لحالها له معنى مستقل  
ماله أي سياق فما نقدر نفهم الكلمة هادي ايش معناها بالضبط  
المعلومات: بس لو حطيت البيانات في سياق context معين فيها structure  
تركيب واضح واستفدت معلومة معينة useful ومرتبته ترتيب معين Organized  
المعرفة: عبارة عن معلومات حطيناها في rule منظمة وفق قواعد محددة ... مثال  
بيانات: أحمد - 18/20 - 20/20 - خالد - 19/20 - 10/20 - 8/20 - رياضيات - علوم  
- اللغة الإنجليزية  
معلومات: درجات احمد في المواد هي كالتالي: رياضيات 20/20-علوم18/20- اللغة  
الإنجليزية-19/20 ---- درجات خالد في المواد هي كالتالي: رياضيات 2/20 - علوم  
8/20-اللغة الإنجليزية 10/20  
المعرفة: احمد ولد شاطر وذكي --- خالد يحتاج اهتمام ومتابعة من المدرسين اكثر  
من احمد  
الحكمة: لو في مسابقات نختار أحمد لانه شاطر --- لازم نهتم في خالد ونعطيه  
دروس خصوصية



# Data, Information, Knowledge , wisdom

	<b>Data</b>	<b>Information</b>
<b>Meaning</b>	Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.	When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.
<b>Example</b>	Each student's test score is one piece of data.	The average score of a class or of the entire school is information that can be derived from the given data

# Data, Information, Knowledge , wisdom

1201	Ahmed	1205	1210	1208
90/100	D	Salim	80/100	
C	Khalid	70/100	B	
Rayan	60/100	A		

These are Data , unorganized facts that need to be processed

# Data, Information, Knowledge , wisdom

ID	Name	Mark	Grade
1201	Rayan	90/100	A
1205	Khalid	60/100	D
1208	Ahmed	70/100	C
1210	Salim	80/100	B

data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.

# What is Data Visualization?

## ماهي البيانات التصورية؟

- ▶ Visualization is the graphic representation of data through the use of pictorial design.
- ▶ هو تمثيل نفس البيانات بطريقة تصويرية لتوضيح البيانات مما يؤدي الى فهمها بشكل اسرع

# Data Visualization Process Explained

1

1

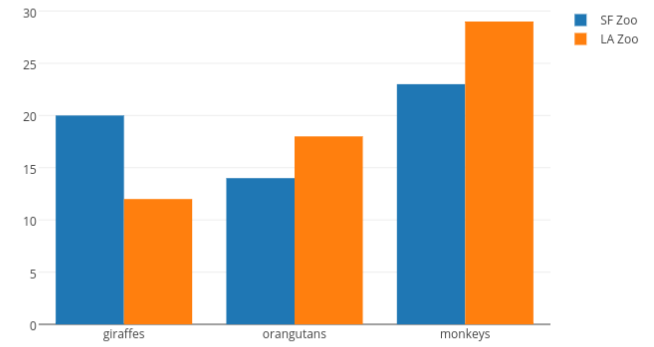
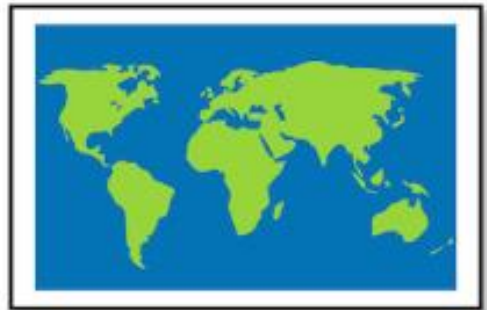
3

3

DATA

INFORMATION

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	Key	Name	UID	Score	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
2	STUDENT A	A	000000001	184	D	A	A	B	D	A	D	B	B	D	B	A	C	A	C
3	STUDENT B	B	000000002	180	D	A	A	B	D	A	D	B	B	D	B	A	C	A	C
4	STUDENT C	C	000000003	156	D	A	A	B	D	A	D	B	C	D	B	A	C	A	C
5	STUDENT D	D	000000004	160	D	A	A	B	D	A	C	B	D	D	A	A	C	A	C
6	STUDENT E	E	000000005	152	D	A	B	D	A	C	B	D	B	D	B	A	C	A	C
7	STUDENT F	F	000000006	164	D	B	A	B	D	A	B	B	D	B	A	C	A	C	
8	STUDENT G	G	000000007	168	D	A	A	B	D	A	C	B	A	D	B	A	C	A	C
9	STUDENT H	H	000000008	136	D	A	A	B	D	A	C	B	B	C	A	A	C	A	C
10	STUDENT I	I	000000009	180	D	A	A	B	D	A	B	B	D	B	A	C	A	C	
11	STUDENT J	J	000000010	180	A	A	B	D	A	D	B	D	B	A	C	A	C		
12	STUDENT K	K	000000011	136	B	C	B	D	A	D	B	C	D	A	A	C	A	C	
13	STUDENT L	L	000000012	132	A	A	B	D	A	C	B	B	D	B	A	C	A	A	
14	STUDENT M	M	000000013	160	D	B	A	B	D	A	D	B	B	D	B	A	C	A	C
15	STUDENT N	N	000000014	136	D	A	A	B	D	A	C	B	B	A	A	C	A	C	
16	STUDENT O	O	000000015	168	D	A	A	B	D	A	C	B	B	C	B	A	C	A	C
17	STUDENT P	P	000000016	168	A	A	B	D	A	D	B	B	D	B	B	A	C	A	C
18	STUDENT Q	Q	000000017	140	D	A	A	B	D	A	A	B	D	B	A	C	A	C	
19	STUDENT R	R	000000018	176	D	A	A	B	D	A	B	B	D	B	B	A	C	A	C
20	STUDENT S	S	000000019	196	D	A	A	B	D	A	D	B	B	D	B	A	C	A	C
21	STUDENT T	T	000000020	136	D	B	C	B	D	A	C	B	C	B	A	C	A	C	



2

4

4



# Data Visualization Process Explained

1



1



3



2



4



4



# Why Data Visualization

- ▶ The goal of a visual is to make information easy to comprehend.
- ▶ الهدف الأساسي من تصوير البيانات هو وصف المعلومة بشكل أوضح مما يؤدي إلى فهمها بشكل أسرع
- ▶ □ Provide qualitative overview of large data sets
- ▶ تعطي نظرة عامة لمجموعة كبيرة من البيانات
- ▶ Search for patterns, trends, structure, irregularities, relationships among
- ▶ Data
- ▶ سهولة البحث عن أنماط وعلاقات بين البيانات
- ▶ □ Help find interesting regions and suitable parameters for further
- ▶ quantitative analysis
- ▶ تساعد في اكتشاف مناطق مثيرة لإهتمام لعمل مزيد من الأبحاث والتحقيق

# Types of Data Visualization


## 1- Infographics



# Data Science Workflow

Business Understanding 

 Data Understanding

 Data Importing

 Data Cleaning & Manipulation

Statistical Modeling & Machine Learning 

Reporting & Visualization 

## The Biggest Shift Since The Industrial Revolution

**Google**  
**You Tube** Youtube is the **second** largest search engine on the web  
**50%** of YouTube's 300 million users go at least once a week.

**Yahoo!**  
**bing**

**3.5 billion** pieces of content are shared each WEEK on facebook.

Facebook usage is up **40%** since last year.

there are over **70 translations** available for facebook.

The average user on facebook has **120 friends**

**65 million** facebook users access the site through their mobile phones. That's up **100%** in the last 6 months.

**2.5 billion** photos are uploaded to facebook each month

**96%** of 18-35 yr olds. are on a social network

**34%** of bloggers post opinions about products & brands. ...

**15%** of bloggers spend 10 hours or more each week blogging.

**78%** of consumers trust peer recommendations.  
**14%** trust advertisements

**25%** of search results for the World's Top 20 largest brands are links to user-generated content.

**1 in 5** americans (age 18-35) use twitter.

So far this year, there are over **30 million** tweets a day.

**80%** of Twitter usage is on mobile devices. Imagine what that means for bad customer experiences???

Twitter has **75,000,000** users, and is adding **300,000** a day!

  
en:gauge media  
www.WeCreateFans.com

# Types of Data Visualization

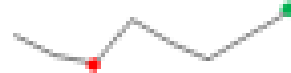
## 2- Sparkline

# Types of Sparklines

Regular Sparkline



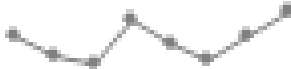
Min and Max Points



First and Last Points



Markers



Column Chart



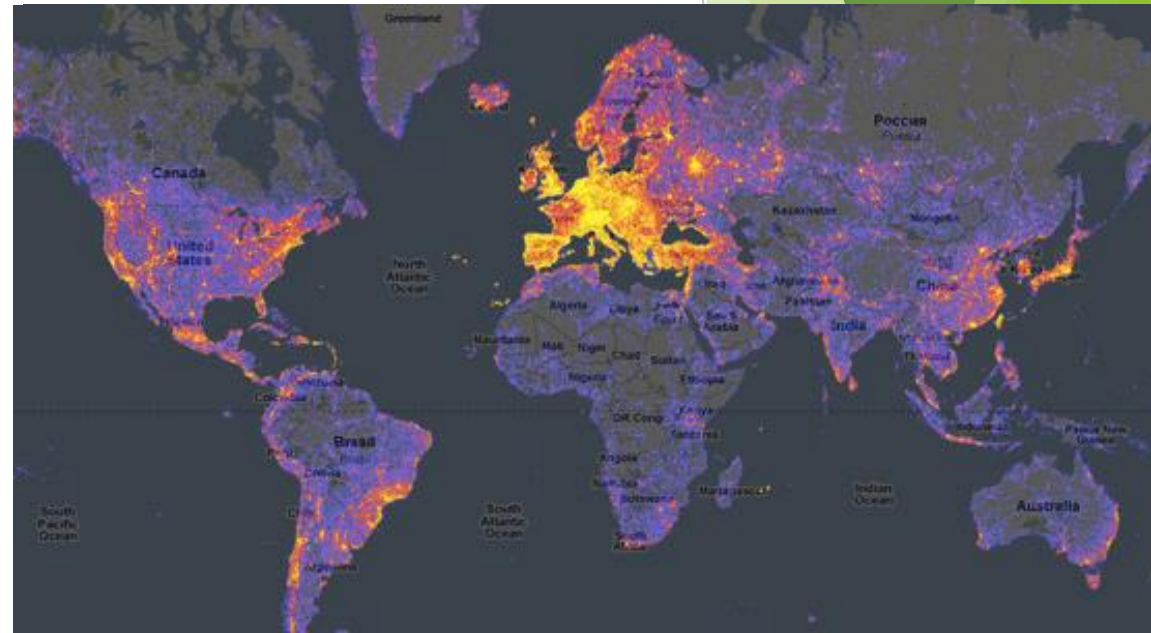
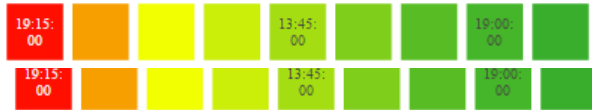
Win Loss Chart



# Types of Data Visualization

## 3- Heatmaps

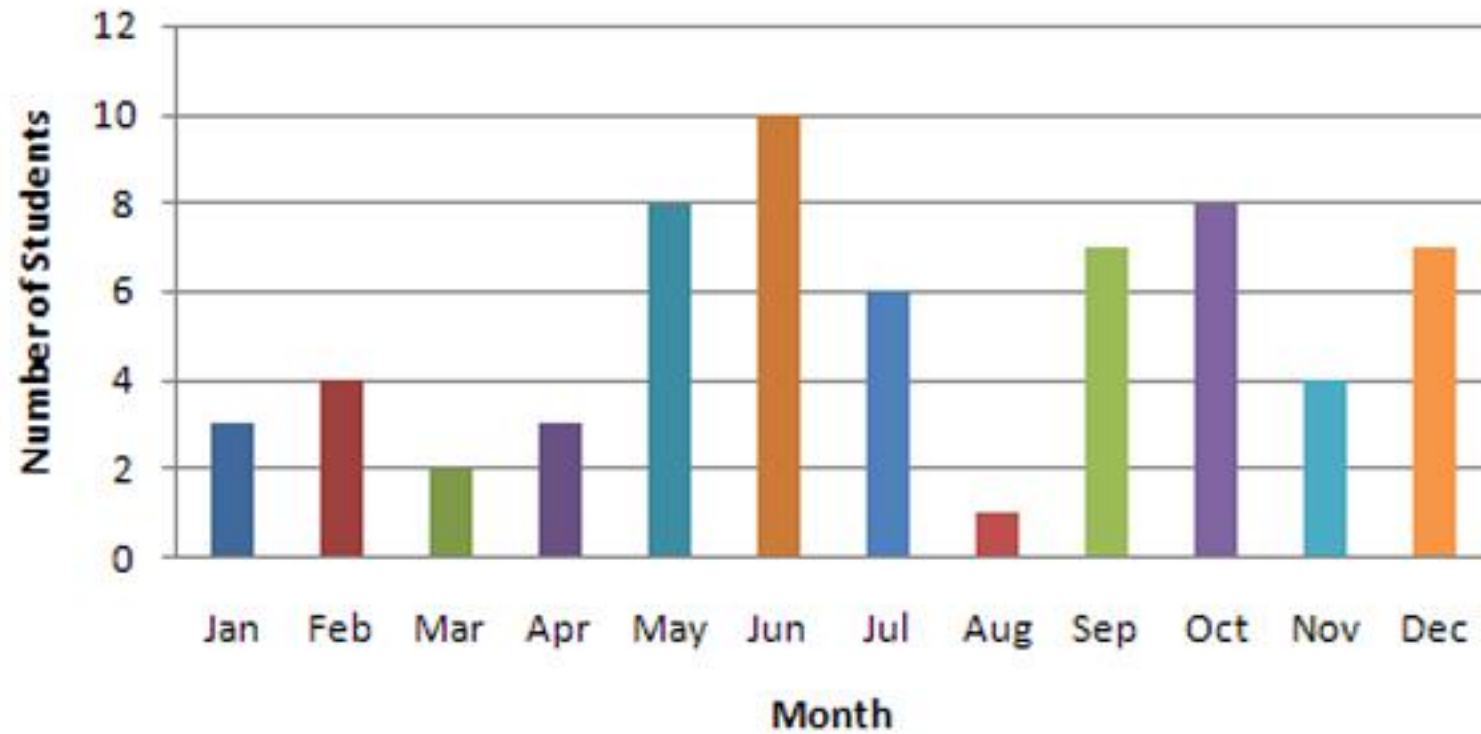
01:00:00	21:45:00	19:30:00	17:15:00	15:00:00	01:45:00	10:30...	10:15...	10:00...	09:45...	09:30...	09:15...	09:00...
01:15:00	21:30:00	19:15:00	17:00:00	14:45:00	23:45:00	08:45:00	07:00:00	06:45:00	06:30:00	06:15:00	06:00:00	05:45:00
23:30:00	21:15:00	19:00:00	16:45:00	14:30:00	12:15:00	08:30:00	08:15:00	05:30...	04:15...	04:00...	03:45...	03:30...
23:15:00	21:00:00	18:45:00	16:30:00	14:15:00	12:00:00	08:15:00	08:00:00	05:15...	03:15:00	02:15:00	02:00:00	01:30:00
23:00:00	20:45:00	18:30:00	16:15:00	14:00:00	11:45:00	07:45:00	07:30:00	05:00...	03:00:00	02:45:00	12:45:00	12:30:00
22:45:00	20:30:00	18:15:00	16:00:00	13:45:00	11:30:00	07:30:00	07:15:00	04:45...	02:30:00	02:15:00	00:00:00	00:00:00
22:30:00	20:15:00	18:00:00	15:45:00	13:30:00	11:15:00	07:15:00	07:00:00	04:30...	02:15:00	02:00:00	00:00:00	00:00:00
22:15:00	20:00:00	17:45:00	15:30:00	13:15:00	11:00:00	07:00:00	06:45:00	04:15:00	02:00:00	01:45:00	00:00:00	00:00:00
22:00:00	19:45:00	17:30:00	15:15:00	13:00:00	10:45:00	06:45:00	06:30:00	04:00:00	01:45:00	01:30:00	00:00:00	00:00:00



# Types of Data Visualization

## 4- BarGraph

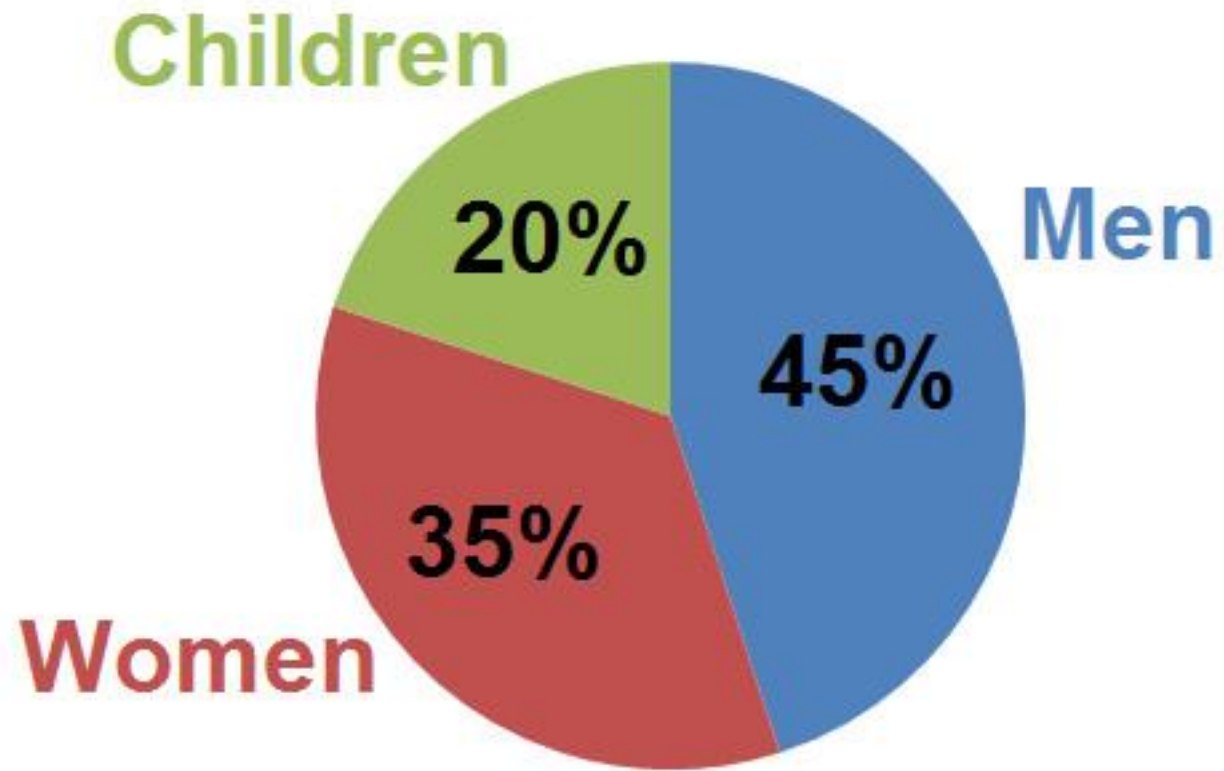
## *Birthday of Students by Month*



# Types of Data Visualization

4- pie graph (or pie chart)

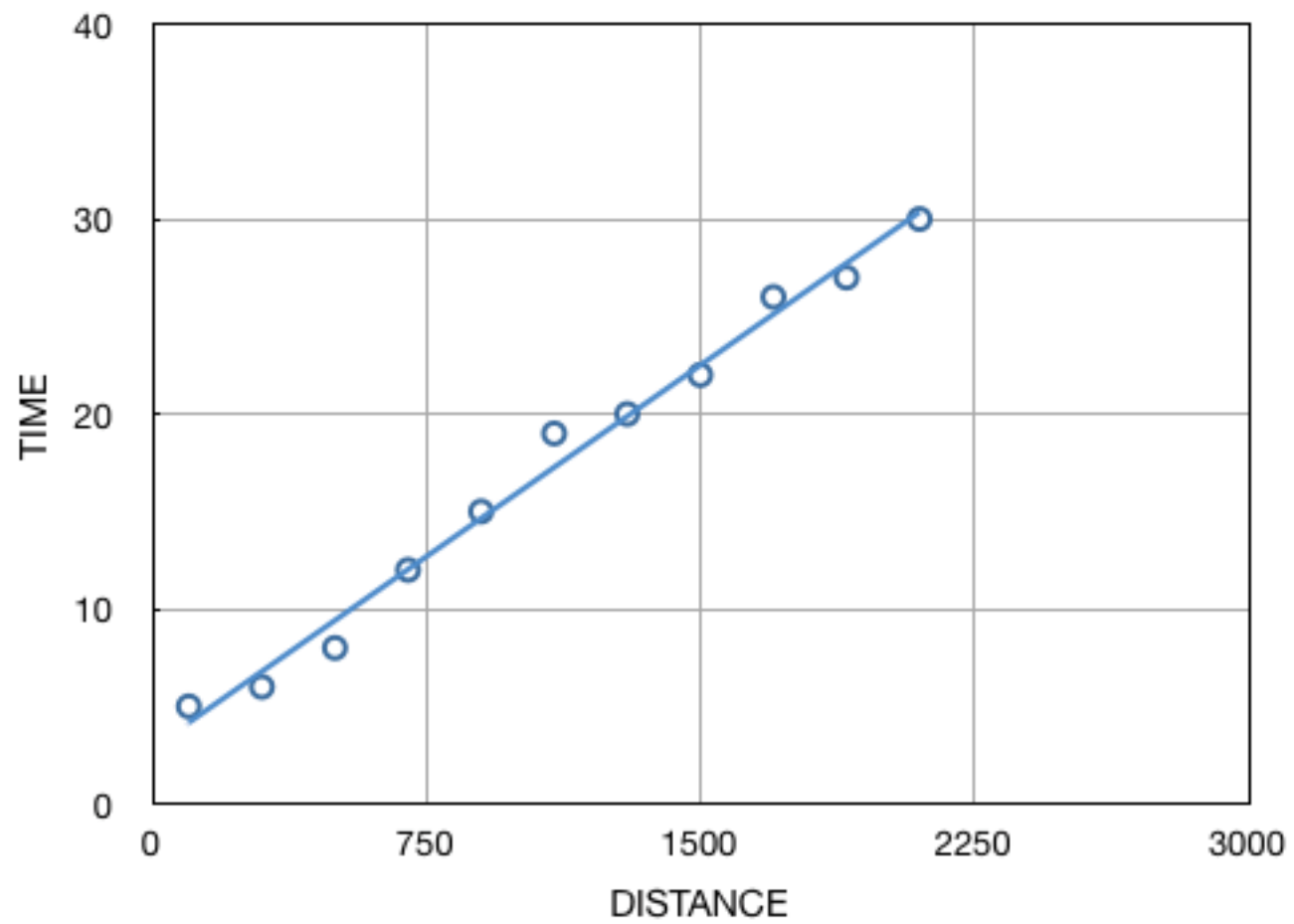




# Types of Data Visualization

## 6- Scatter plot

Scatter Plot 1





Google  
Analytics

Google Analytics is a freemium web analytics service offered by Google that tracks and reports website traffic

Search reports & help

All Users  
100.00% Sessions

Mobile and Tablet Traffic  
16.02% Sessions

+

Overview

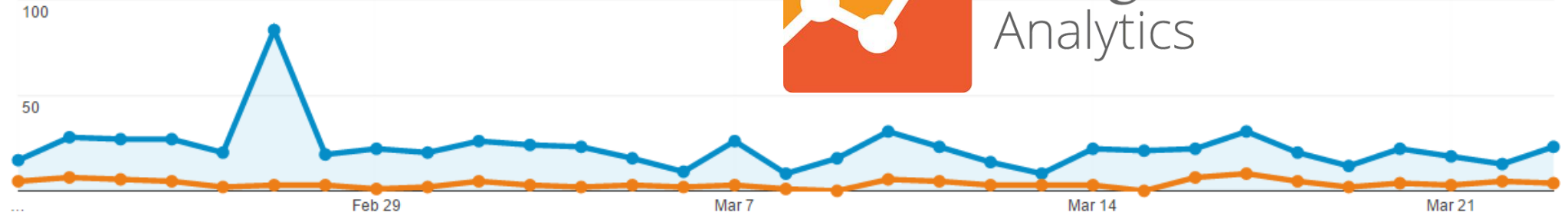
Sessions vs. Select a metric

Hourly Day Week Month

Sessions (All Users) Sessions (Mobile and Tablet Traffic)



Google Analytics



Sessions  
All Users  
**699**

Mobile and Tablet Traffic  
**112**

Users  
All Users  
**627**

Mobile and Tablet Traffic  
**99**

Pageviews  
All Users  
**1,217**

Mobile and Tablet Traffic  
**188**

Pages / Session  
All Users  
**1.74**

Mobile and Tablet Traffic

Avg. Session Duration  
All Users  
**00:01:07**

Mobile and Tablet Traffic

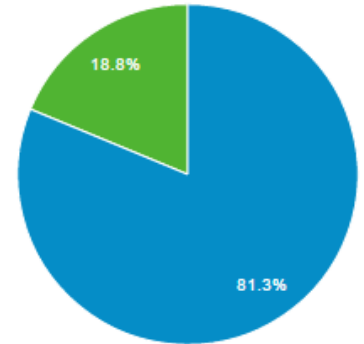
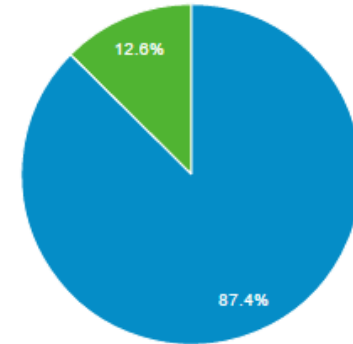
Bounce Rate  
All Users  
**73.68%**

Mobile and Tablet Traffic

New Visitor Returning Visitor

All Users

Mobile and Tablet Traffic



Dashboards

Shortcuts

Intelligence Events

Real-Time

Audience

Overview

Active Users

Cohort Analysis **BETA**

Demographics

Interests

Geo

Behavior

Technology

Mobile

Custom

Benchmarking



Google Data Studio

**Google Data Studio** (beta) turns your data into informative dashboards and reports that are easy to read, easy to share, and fully customizable

iOS Sessions  
**56,770**

Android Sessions  
**49,799**

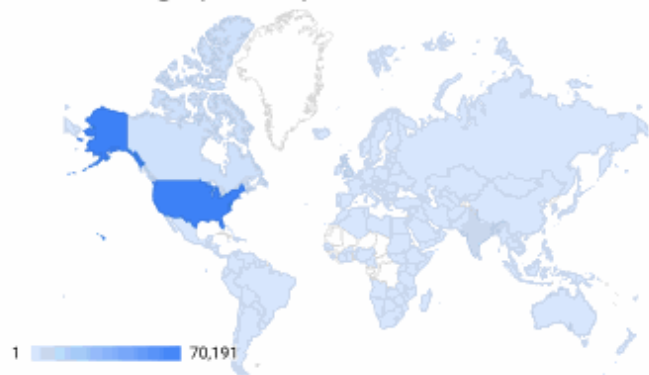
Other Sessions  
**3,829**



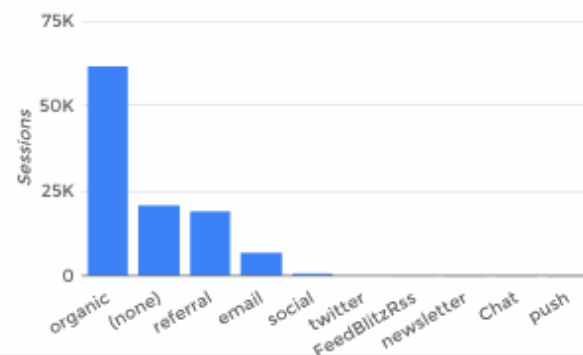
## Mobile Operating System Breakdown

Operating System	Users	Sessions	% New Sessions	Bounce Rate	Avg. Session Duration
1. iOS	43,416	56,770	67.74%	89.05%	00:00:55
2. Android	38,831	49,799	71.14%	85.08%	00:01:10
3. (not set)	853	1,057	73.04%	79.47%	00:01:36
4. Windows Phone	843	1,056	68.66%	85.13%	00:01:06
5. Windows	788	939	72.20%	82.00%	00:01:29
6. BlackBerry	498	622	68.81%	86.01%	00:01:11
7. Nokia	45	58	74.14%	84.48%	00:00:59
8. Samsung	36	41	80.49%	85.37%	00:00:42
9. SymbianOS	37	41	90.24%	85.37%	00:01:16
10. Playstation Vita	8	8	100.00%	87.50%	00:01:04

## Mobile Geographic Report



## Mobile Sessions by Medium



## Mobile Goal Analysis

Goal Completions  
**7,439**

Goal Conversion Rate  
**6.74%**





# QlikView

QlikView is the most flexible Business Intelligence platform for turning data into knowledge. More than 24,000 organisations worldwide have enabled their users to easily consolidate, search, and visually analyse all their data for unprecedented business insight using QlikView's simplicity.





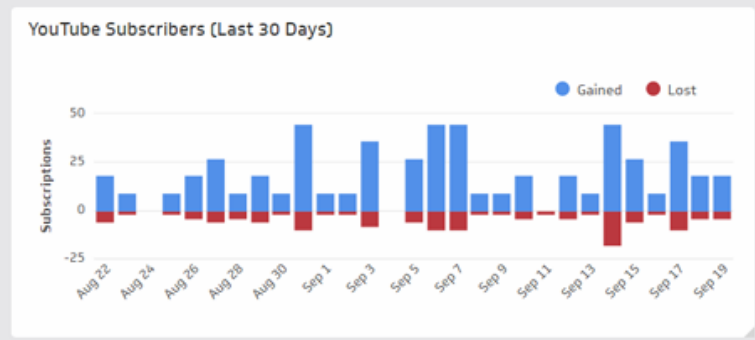
# QlikView

### Social Media Followers

**26,472**  
Likes

**45,322**  
Followers

**6,524**  
Followers



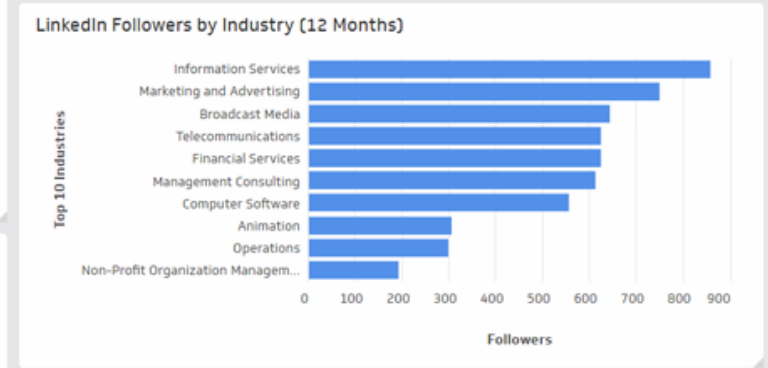
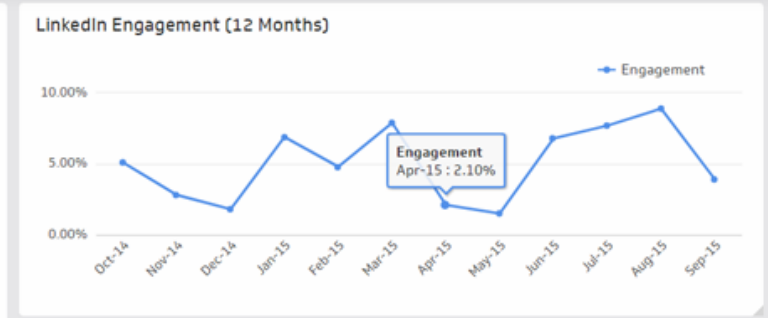
### Twitter Stats (@:)

**5,942**  
Followers

**4,709**  
Friends

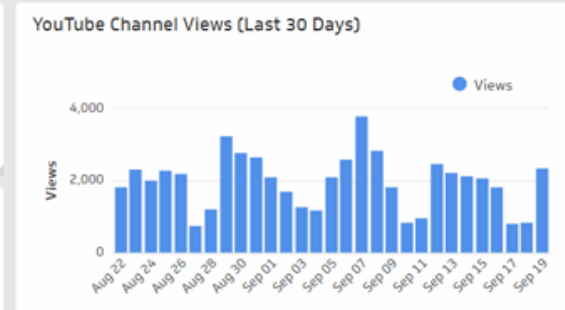
**9,633**  
Statuses

**5,218**  
Favourites



### Twitter Retweets of Last 10 Tweets (@:)

**418** Retweets ▲ 141  
vs 277 (prev.)



### Google Analytics Social Media Conversions

Apr 01, 2015 to Apr 30, 2015 | Last 30 Days

Platform	fb	b4	Conversion %
Twitter	1,964	-158	4%
LinkedIn	26	10	4%
Facebook	14	6	2%
Quora	12	-14	2%
Google+	4	-6	2%
Blogger	2	2	10%
Pocket	2	2	7%
Youtube	2	0	17%

### Facebook Page Stats

**872**  
People talking about your page

**3,076**  
New likes of your page

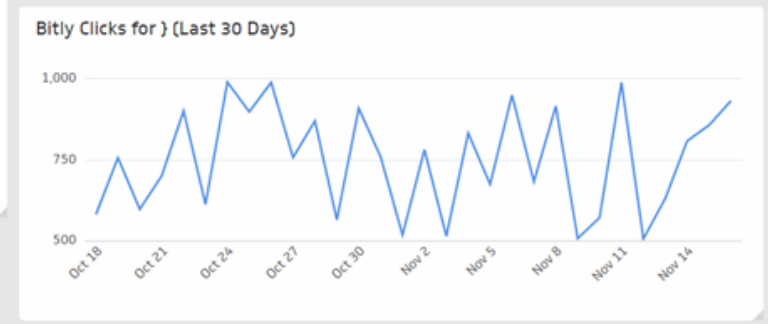


### LinkedIn Engagement

**1.97%** Engagement last month ▲ 5%  
vs 1.87% prev. month

### Facebook Page Stats

**8,721**  
Total likes of your page





Information Builders is a leader in Business Intelligence, Data Integration and Data Quality solutions. Their products and services have received top recognition from independent analyst research firms including Gartner, Forrester, Ventana Research, BARC, Dresner, and the Data Warehouse Institute

Predictive Readmissions Performance

**Selection Criteria**

Chronic Disease: ALL ANEMIA COPD DIABETES HYPERTENSION OTHER

Insurance: ALL A123 B234 C345 D456

Age Range: ALL 20 - 44 45 - 64 65+

Ethnicity: ALL A B C D

Readmit Likelihood >: 14.33%

Mortality Risk: H

Select Sort Values: Age Range Age Ethnicity Insurance Chronic Disease

Gender

Output Formats: HTML PDF Excel Active HTML Active PDF

Update Deferred

Patients Predicted to Readmit: 166

Total Patients: 249

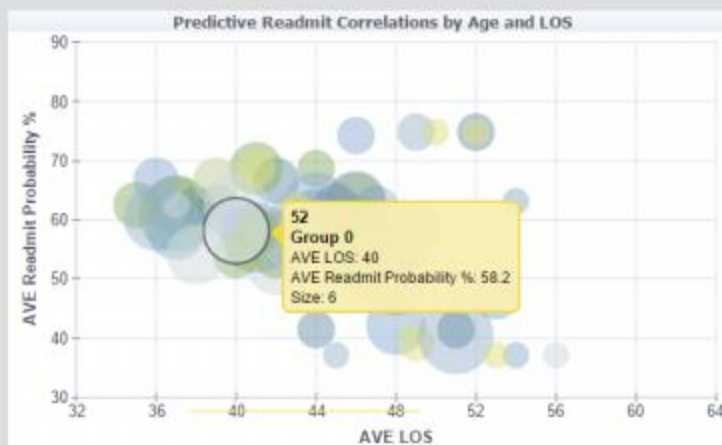
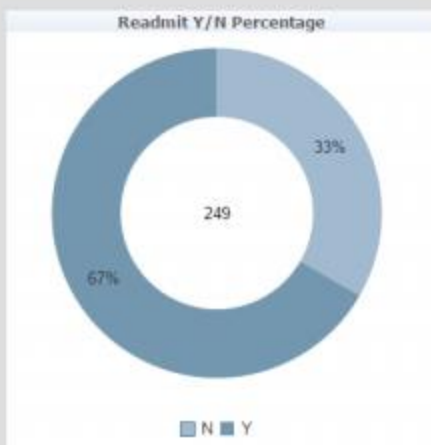
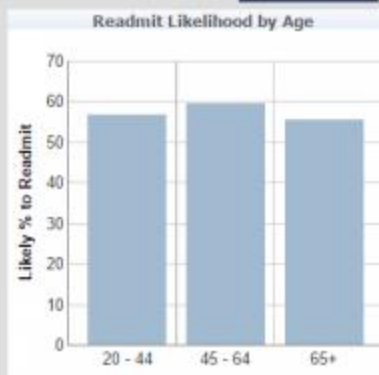
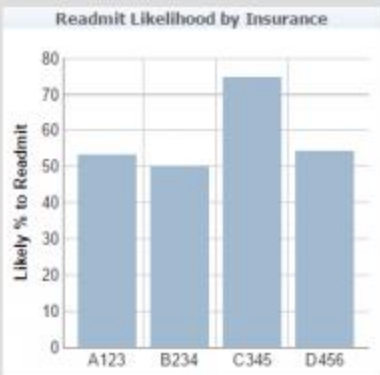
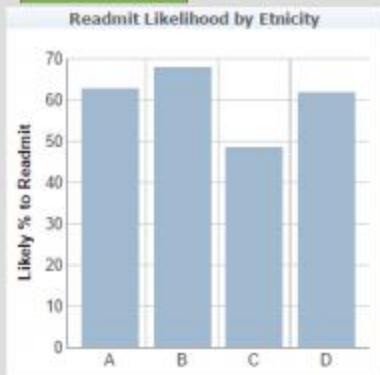
Patients with High Mortality Risk: 249

ALOS: 43

Patients with Overdue Appts: 53

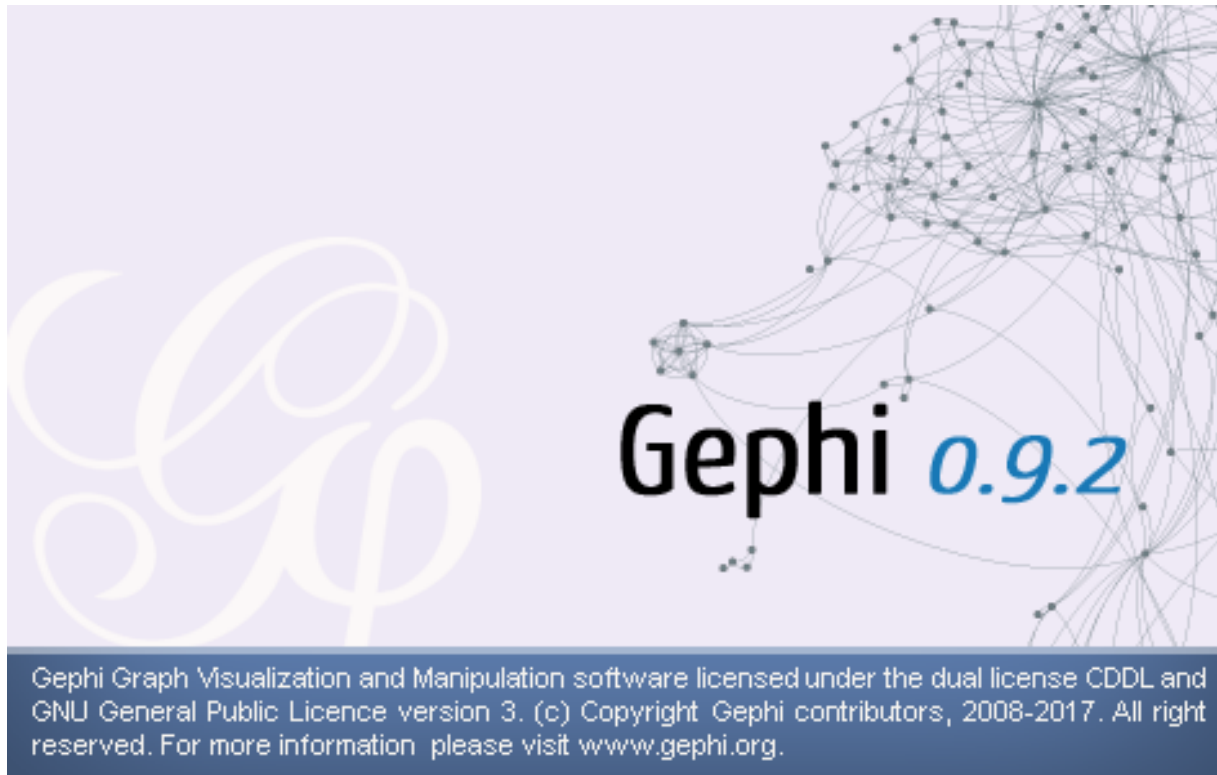
Performance

Patient Detail



Patient Work List

Patient	Likely to Readmit	Readmit Prob. %	Next P/C Pr
1 Y	62.50	62.50	
2 Y	62.50	62.50	
3 Y	74.84	74.84	
4 Y	62.50	62.50	
5 Y	62.50	62.50	
6 Y	62.50	62.50	
7 Y	62.50	62.50	
29 Y	74.84	74.84	
30 Y	62.50	62.50	
31 Y	62.50	62.50	
32 Y	62.50	62.50	
33 Y	62.50	62.50	
34 Y	62.50	62.50	
35 Y	62.50	62.50	
57 N	37.11	37.11	
58 Y	62.50	62.50	
59 Y	62.50%	62.50%	03/13/2014
60 Y	62.50%	62.50%	03/13/2014
61 Y	62.50%	62.50%	04/10/2014
62 Y	62.50%	62.50%	04/10/2014
63 Y	74.84%	74.84%	04/10/2014
85 Y	62.50%	62.50%	05/04/2011
86 Y	74.84%	74.84%	05/04/2011
87 Y	74.84%	74.84%	05/04/2011
88 Y	74.84%	74.84%	05/04/2011
89 Y	74.84%	74.84%	05/04/2011
90 Y	62.50%	62.50%	05/04/2011
91 Y	62.50%	62.50%	07/18/2011
113 Y	62.50%	62.50%	07/18/2011
114 Y	62.50%	62.50%	07/18/2011
115 N	41.44%	41.44%	07/18/2011
116 Y	62.50%	62.50%	07/18/2011
117 Y	62.50%	62.50%	0
118 N	41.44%	41.44%	0
119 Y	62.50%	62.50%	0



**Gephi** is an open-source software for network visualization and analysis. It helps data analysts to intuitively reveal patterns and trends, highlight outliers and tells stories with their data. It uses a 3D render engine to display large graphs in real-time and to speed up the exploration.



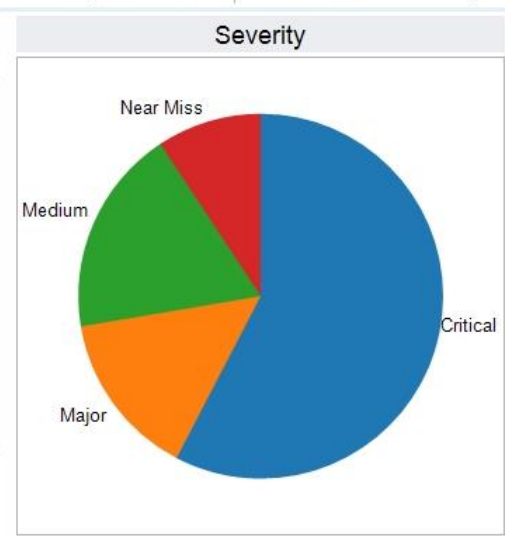
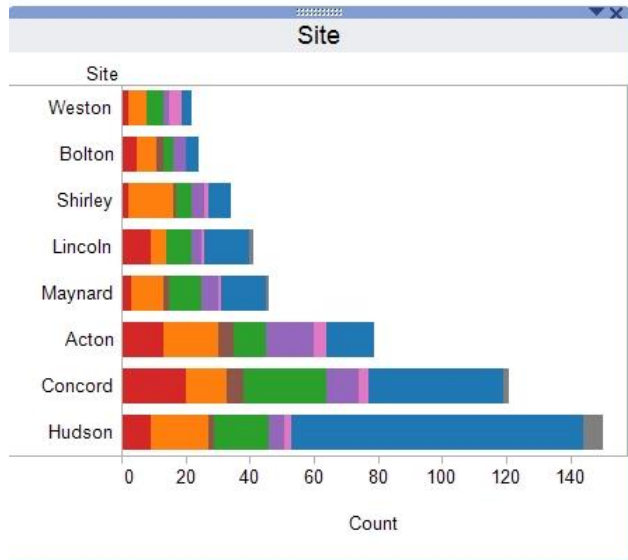
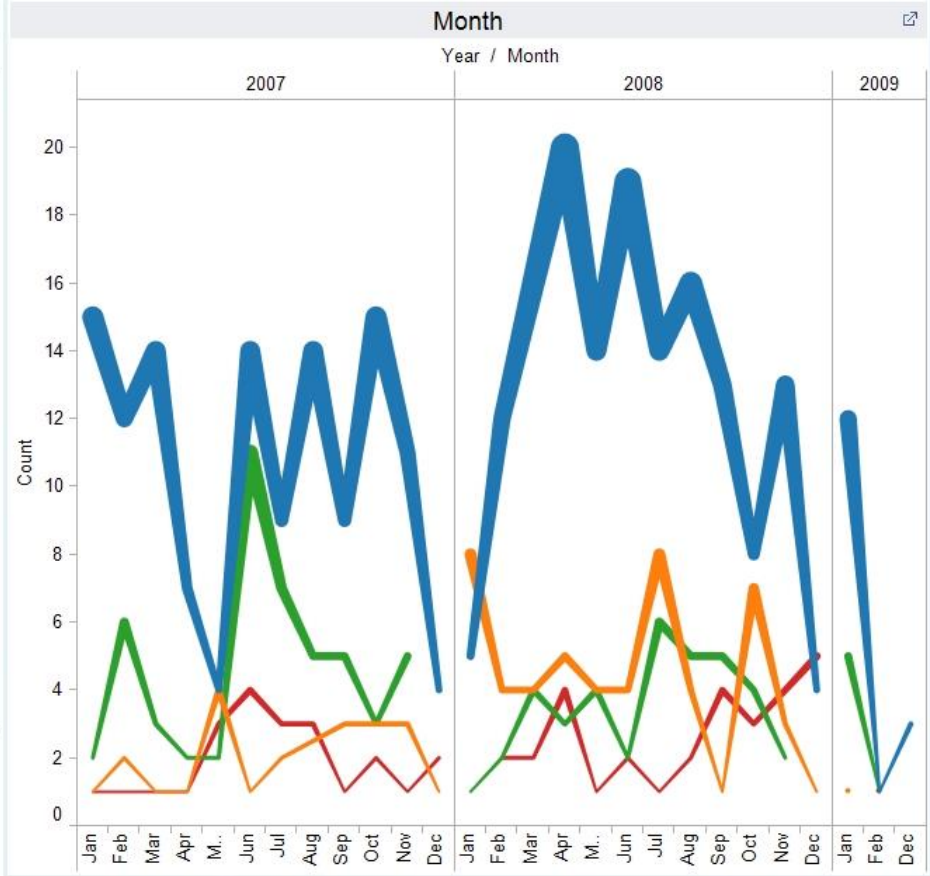
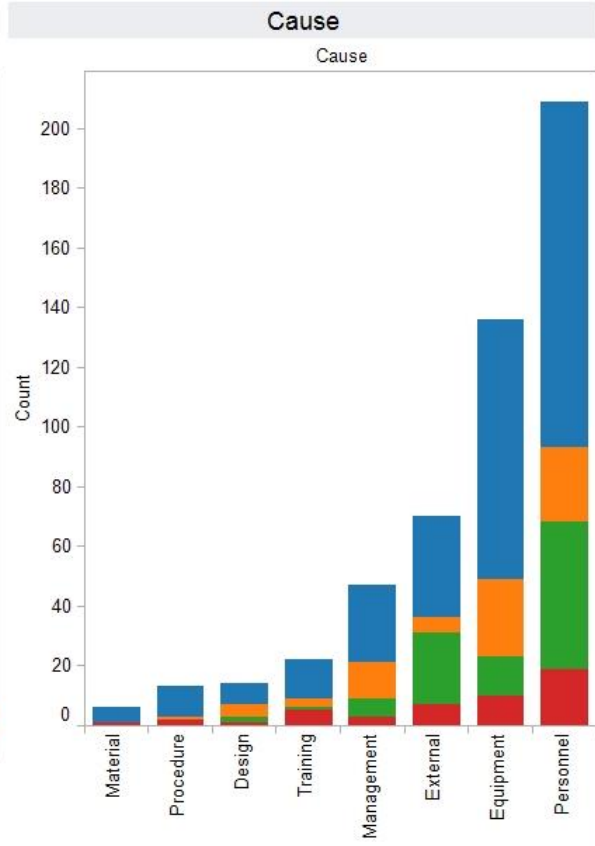
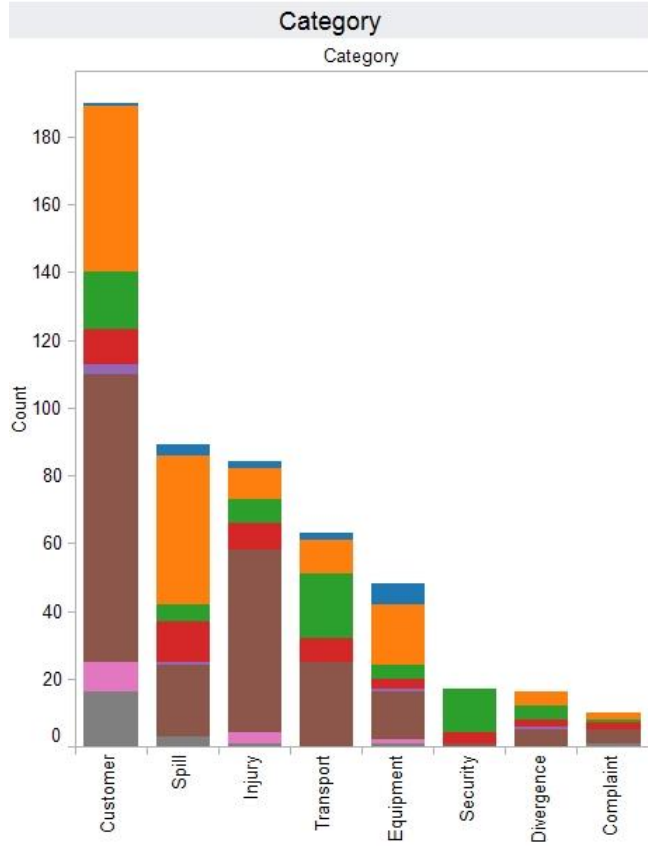


**Tableau** is an interactive data visualization tool that can be used by even non-technical professionals to come up with customized dashboards and generate compelling business insights using a wide variety of data sources

In 2003 Tableau spun out of Stanford University with VizQL™, a technology that completely changes working with data by allowing simple drag and drop functions to create sophisticated visualizations. The fundamental innovation is a patented query language that translates your actions into a database query and then expresses the response graphically.

Read more at

<https://www.tableau.com/products/technology#oVAYmFtJyH3Mo5Wq.99>



- Category**
  - (All)
  - Complaint
  - Customer
  - Divergence
  - Equipment
  - Injury
  - Security
  - Spill
  - Transport
- Site**
  - (All)
  - Acton
  - Bolton
  - Concord
  - Hudson
  - Lincoln
  - Maynard
  - Shirley
  - Weston
- Month**
  - (All)
  - Jan
  - Feb
  - Mar
  - Apr
  - May
  - Jun
  - Jul
  - Aug
  - Sep
  - Oct
  - Nov
  - Dec
- Year**
  - (All)
  - 2007
  - 2008
  - 2009
- Status**
  - (All)
  - Closed
  - Open

**Count**  
517.0



**The End**