

# FUNDAMENTALS OF DATAFLOW AND SQL

*Explain where SQL fits in the dataflow*

*Retrieve and filter data with basic SQL*

*Navigate a Relational Database*

***SELECT***

***FROM***

***WHERE***

***ORDERBY***

***LIMIT***

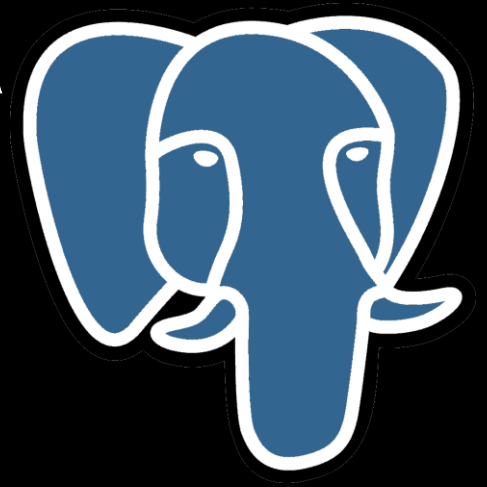
***Matthew Morris***

*Git: Morrisdata*

*MatthewMorris.DA@gmail.com*



# Previously in Data Analytics



*AirBnB presentations*  
*Config and install of ;*  
*Postgres and pgAdmin*



# Prompt for Project Mozilla Firefox

It's 2010 you are attempted to help Mozilla Firefoxes Product Manager decide if they should be investing development efforts in a better Bookmark experience or Tab experience.

Your putting together a presentation recommending an area of focus for the next version of Firefox.

You luckily stumble across a [recent Test Flight](#) with customer surveys and usage analytics to better inform your thinking before deciding which of the two areas to recommend.

# Prompt for Project Mozilla Firefox

## Initial Draft of Findings:

Describe the users represented in the data.

Describe any initial findings related to your research questions.

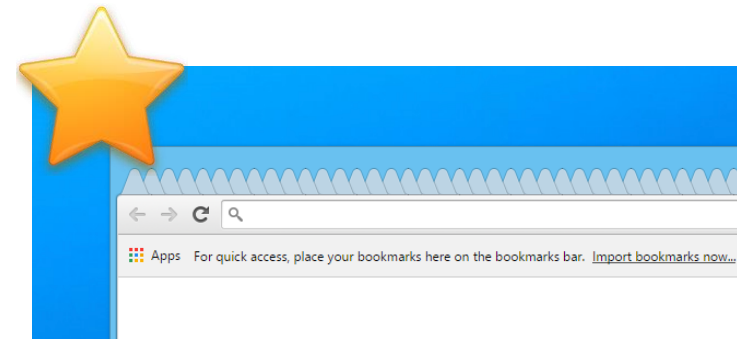
Describe what you still need to do for your presentation in order to make a sound recommendation

## SQL Queries + Questions:

List the appropriate business questions you were trying to answer

Use correct, efficient syntax for the functions and operators needed

Include corresponding output listed in the .sql file



# Prompt for Project Mozilla Firefox

## **10-minute Presentation:**

Summarize the data in your presentation

Summarize your exploration of the bookmark usage feature in your presentation.

Summarize your exploration of usage of browser tabs in presentation.

Compare the two features and create a coherent recommendation to your boss.

Include analysis of data with statistics and outliers

Include recommendation with predicted outcomes and identified action items

Present limitations, assumptions

Identify follow-up problems and questions for future analysis

---

FUNDAMENTALS OF DATAFLOW AND SQL

---

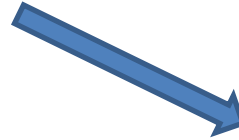
# GATHERING REQUIREMENTS

# GATHERING REQUIREMENTS



***FUNDAMENTALS OF DATAFLOW AND SQL***

# GATHERING REQUIREMENTS



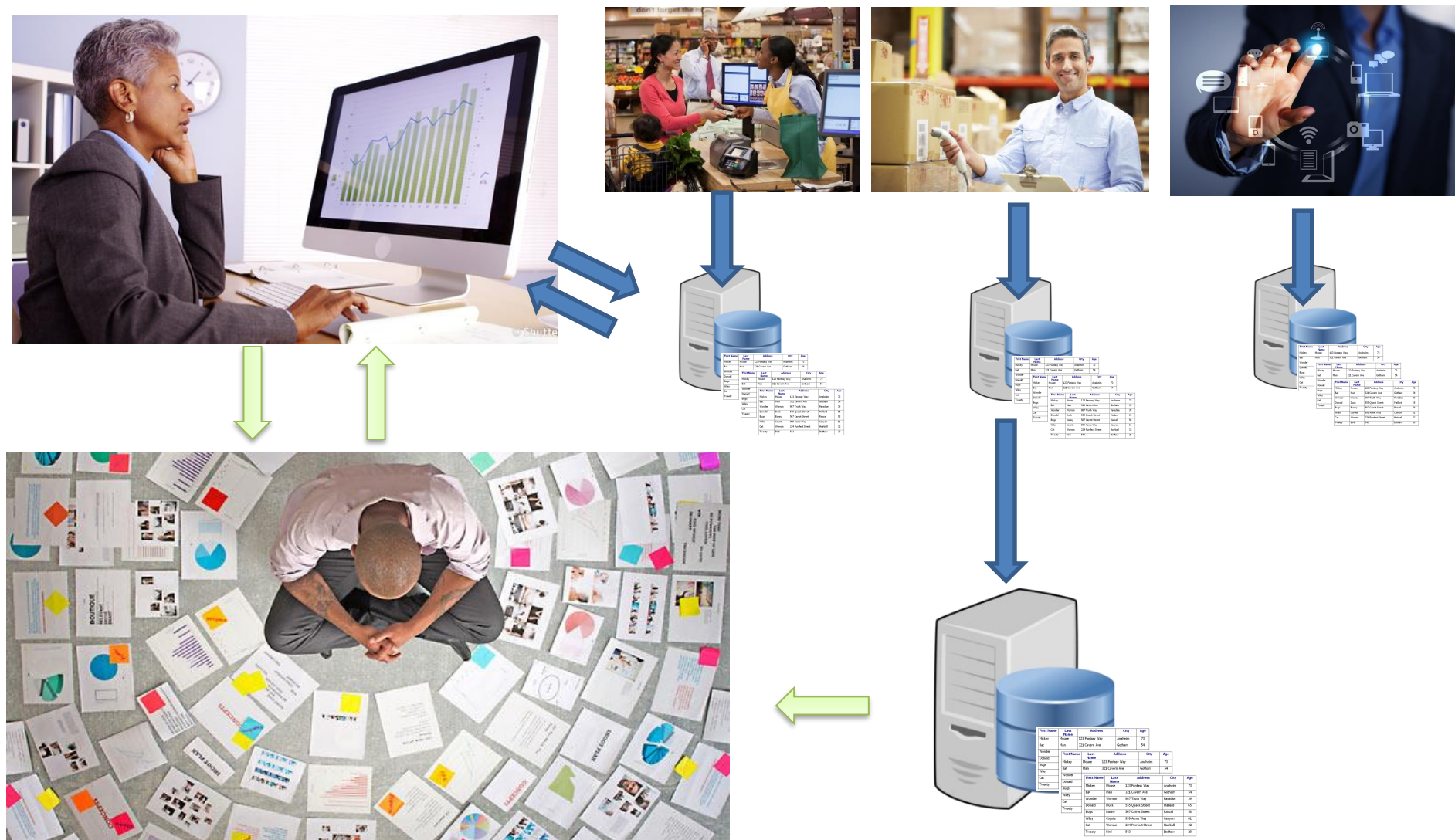
	First Name	Last Name	Address	City	Age
Hickey					
Bat					
Wong	Hickey				
Dona	Bat				
Bugs	Wong	Mouse	123 Fantasy Way	Anaheim	73
Wiley	Dona	Man	321 Cavern Ave	Gotham	54
Cat	Bugs	Wonder	987 Truth Way	Paradise	39
Wiley	Donald	Duck	555 Quack Street	Hallard	65
Cat	Bugs	Bunny	567 Carrot Street	Rascal	58
Tweety	Wiley	Coyote	999 Acme Way	Canyon	61
Cat	Woman		234 Purrfect Street	Hairball	32
Tweety	Bird	543		Ittobrav	28



## FUNDAMENTALS OF DATAFLOW AND SQL



# GATHERING REQUIREMENTS



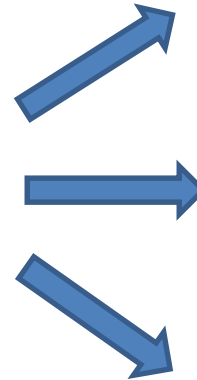
**FUNDAMENTALS OF DATAFLOW AND SQL**

# GATHERING REQUIREMENTS

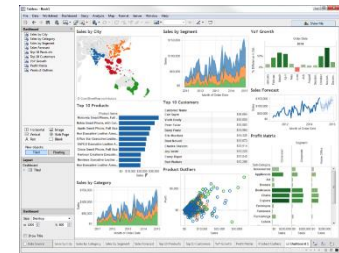


***FUNDAMENTALS OF DATAFLOW AND SQL***

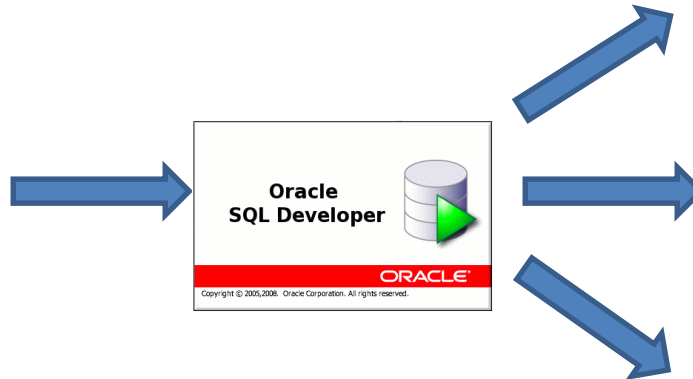
# Basic Ecosystem



**COGNOS**



# Basic Ecosystem

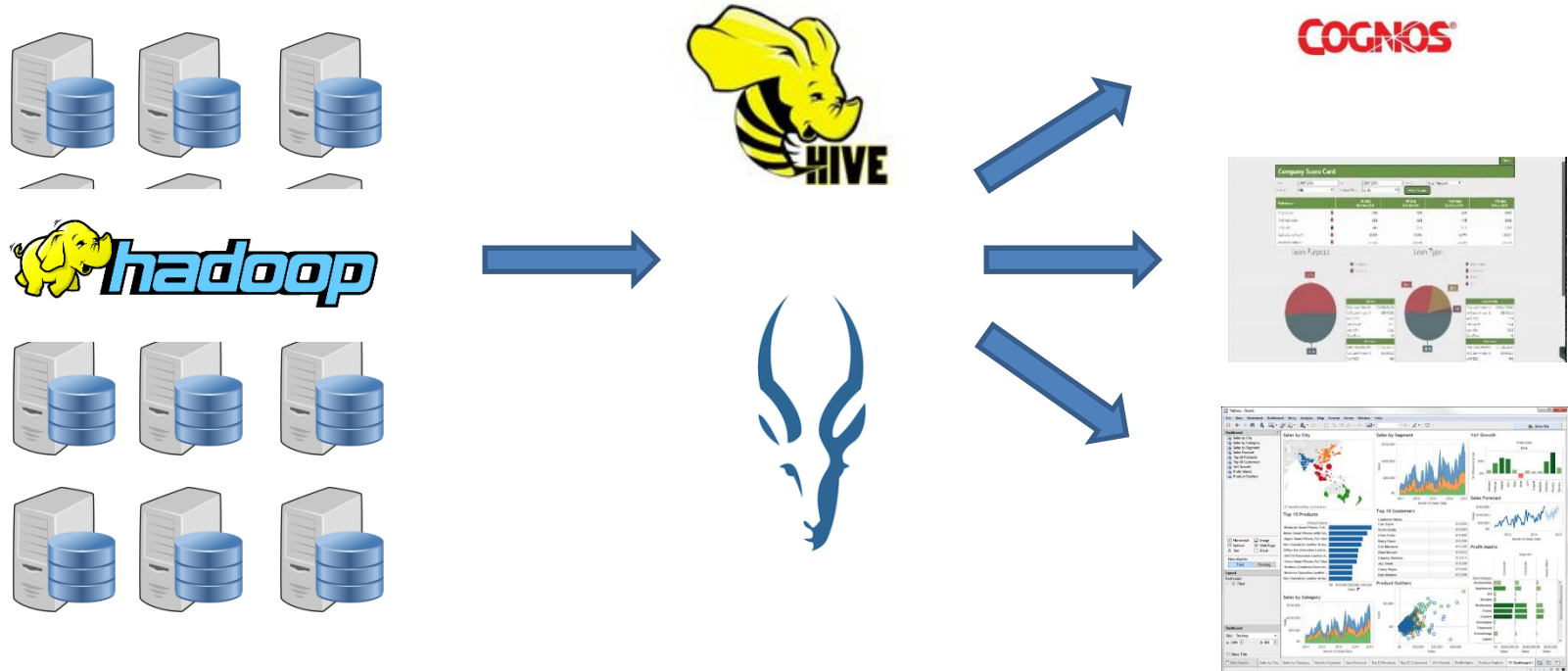


COGNOS®





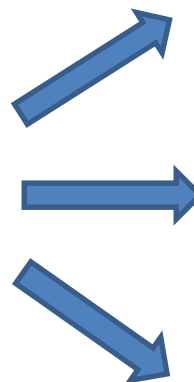
# Basic Ecosystem



# Basic Ecosystem



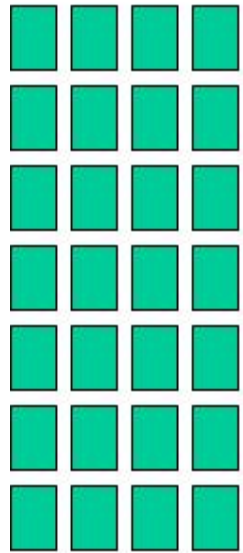
alteryx



COGNOS



Power BI



Libraries/Collections

Schema

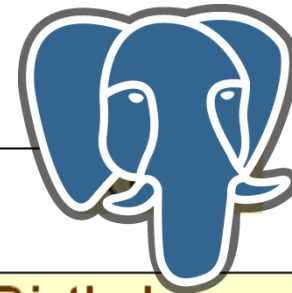
Tables/Files/Objects

Members/Partitions in files



## NETWORK

<u>Contacts</u>	<u>Gift Ideas</u>	<u>Party</u>
Name	Name	Name
Friends	Gift ideas	Holidays
Family	Gifts received	Birthday
Business Name	Gifts given	Anniversaries
Address		
Phone number		
Occupation		



### Birthdays

Name  
Birthday  
Gift Ideas  
Gifts received  
Gifts given  
Family



Understanding your data

What are the tables

What are the fields

How might you link the  
tables?



SELECT  
FROM  
WHERE  
ORDER BY  
LIMIT

SELECT \*

SELECT FIELD1, FIELD2 ...

SELECT (FIELD1+FIELD2), FIELD 3...

SELECT SUM(FIELD1), FIELD2

```
SELECT DISTINCT Location, NumberOfSales
```

Location	NumberOfSales
Seattle	101
Seattle	40
Tacoma	72

```
SELECT DISTINCT Location, NumberOfSales, Date
```

Location	NumberOfSales	Date
Seattle	101	10/28/17
Seattle	101	10/27/17
Seattle	40	10/26/17
Tacoma	72	10/28/17
Tacoma	72	10/27/17

WHERE COUNTRY = US

WHERE COUNTRY = US  
AND STATE = WA

WHERE COUNTRY = US  
AND STATE = WA  
AND SALES > 100

ORDER BY 1

ORDER BY 1,2 DESC

LIMIT 1000

ROWNUM <= 1000



CAST( field AS type)

A helpful String function

1. Select various fields from the SALES table that interest you.  
\*BE sure to use LIMIT 1000
2. Practice using filters.
3. Use AND to apply multiple filters  
Change the sort.
4. Save your Query

Use your new skills to review Mozilla  
Firefox



# Q & A

---

*If we have data, let's look at data.  
If all we have are opinions,  
let's go with mine.*

-Jim Barksdale, Prior CEO Netscape

# Conclusion

SQL is the language of questions and retrieving answers.

Developing good questions is the first step in developing a good SQL Statement.

The next step is knowing where your data resides. In order to find that out you need to find your source behind many reports



# EXIT TICKET

**CLASS : FUNDAMENTALS OF DATA FLOW AND SQL**

**QUESTION:**

**Which Clause would you use to create your Filters?**

