



Leadership

ANALYTICS PROGRAMMING - R

Naveen Kumar

Week 1

Agenda

- Welcome and Introduction
- Course Overview
- Guest Speakers
- Introduction to Data Analytics
- Tools for Data Analysis
- Why R?
- Introduction to R GUI
 - Basic Operations in R GUI
- Summary and Conclusion

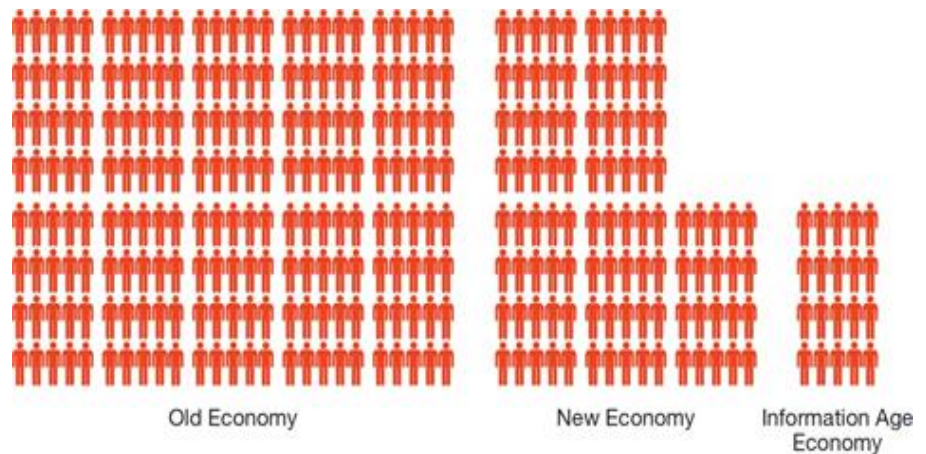
Data: The Root and Purpose of Information Systems

- Alone, raw data are not very useful
- When processed, data transforms into information
- When information is understood and used for decisions, it becomes knowledge

Data	Information	Knowledge
465889727	465-88-9727	465-88-9727 → John Doe
Raw Symbols	Formatted Data	Data Relationships
Meaning: ----- ???	Meaning: ----- SSN	Meaning: ----- SSN → Unique Person

Big Data

- IDC estimates that in 2018, 33 zettabytes of data were generated and consumed
- What would that amount equal? It is 33 trillion gigabytes (IDC, 2018)
- Increasing amounts of data increases the ability to detect meaningful relationships and other insights which can contribute to business success
- Forecasted to grow to 175 zettabytes by 2025

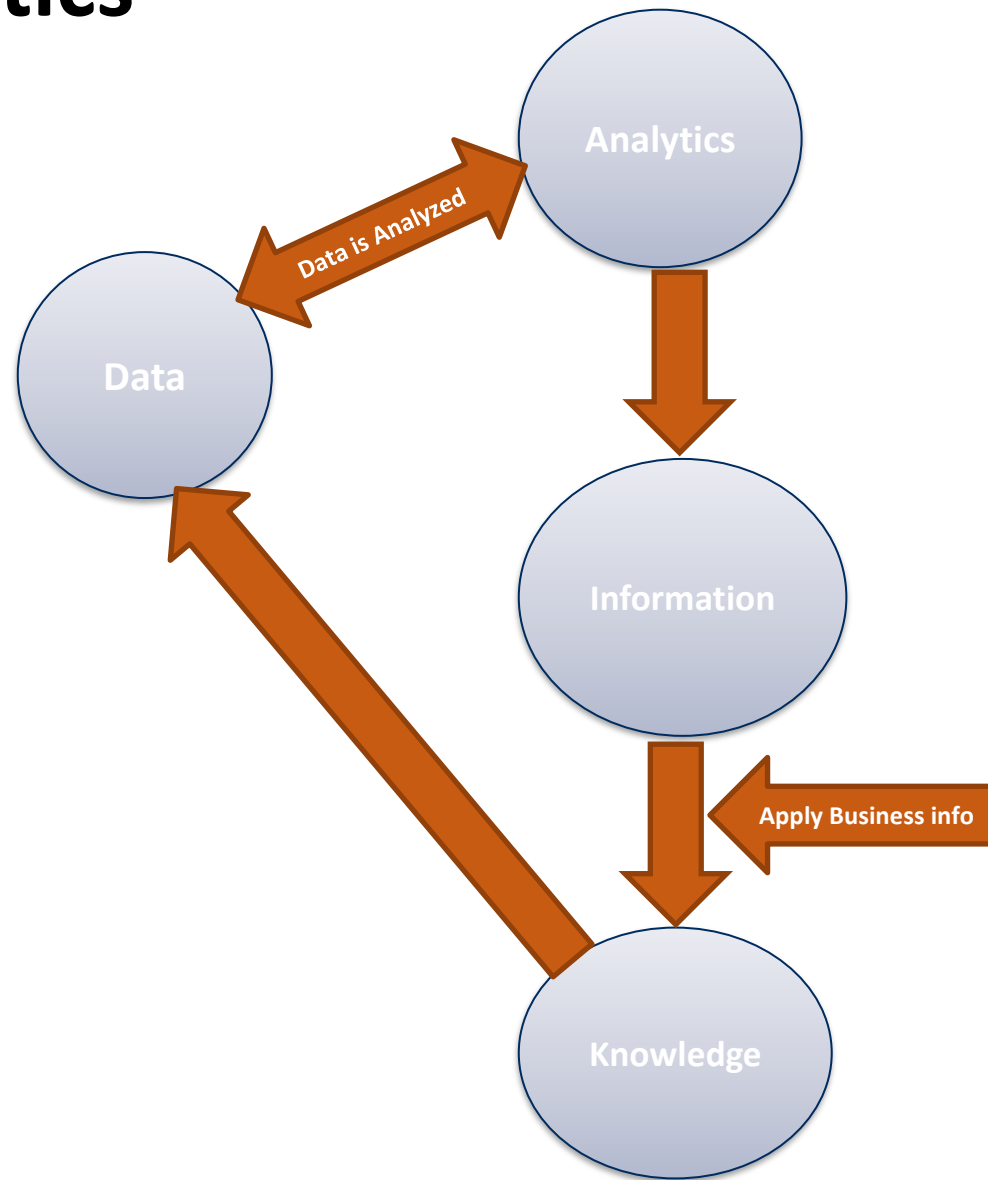


Big Data

- Big Data:
 - High Volume (Lots of it)
 - High Velocity (Accrues quickly)
 - High Variety (Different kinds)
- New technologies and techniques required to capture, store, and analyze big data

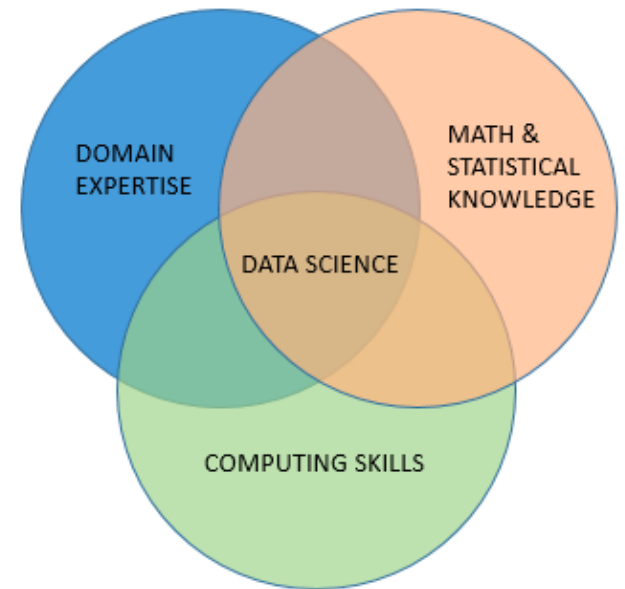
Data Analytics

- Data Analytics
 - Helps to derive meaningful insights or information and subsequently knowledge from data and make critical business decisions
 - Data can be in various forms, either structured or unstructured



Analytics Problem Solving Skills

- Combining concepts or knowledge:
 - Quantitative Skills (Statistics, Math, Operations Research etc.
 - Computer Science Skills (Algorithm Development, Programing, Database etc.)
 - Business/Domain Expertise



Data Analytics Tools

- Data analysis tools
 - Make it easier to process and manipulate data
 - Analyze relationships
 - Identify patterns and trends
 - Build predictive and prescriptive models
- Choosing the right tool is always challenging
- Two Types of tools are available in market
 - Open Source
 - Paid

Open Source and Paid Tools

Open Source Tools	Commercial Tools (Paid)
R	SAS
Python	Tableau
Apache Spark	Excel
Apache Storm	QlikView
PIG and HIVE	Splunk

What is R?

- Programming language
- Produces high quality analysis and graphics
- Open-source
- Free
- Runs on any operating system
- User types commands directly into user “console”

Downloading Cran R

- For Windows

<https://cran.r-project.org/bin/windows/>

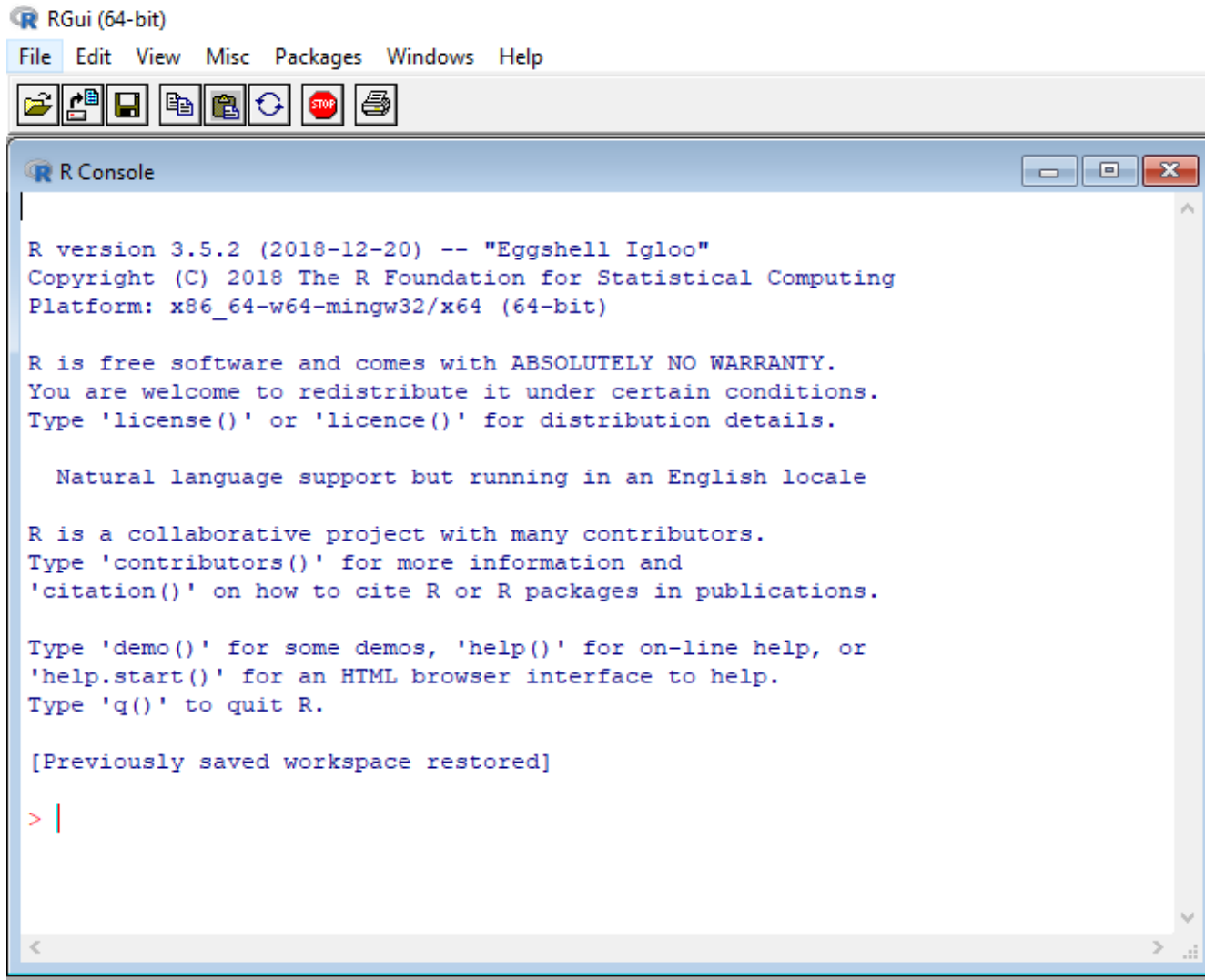
- For Mac

<https://cran.r-project.org/bin/macosx/>

- For Linux

<https://cran.r-project.org/bin/linux/>

R GUI



Introduction to R

- **Typing into R**
 - The > shows where to type into R console window.
 - Commands can spread over more than one line.
 - If cursor displays + then command is not complete (some additional typing is required).
 - Spaces are usually ignored.

Introduction to R

- GUI = Graphical User Interface
- It can do math
 - Add, subtract, multiply, divide, exponents
- Built-in functions
 - Mean, sum, sequence
 - Can also do True/False
- Comments are written starting with #

Key Information

- Version
- Platform
- License
- Contributors
- Citation
- Demo
- Help
- Quit

Basic Operations

- The three basic types of operations used in R GUI (also in R Studio) are:
 - Arithmetic operations
 - Relational operations
 - Logical operations

Arithmetic Operations

- In basic form, R GUI can be used as a simple calculator:
 - Addition (+)
 - Subtraction (-)
 - Multiplication (*)
 - Division (/)
 - Exponentiation (^ / **)
 - Exponentiation is used for taking the power
 - Modulo (%)
 - Returns remainder of the division of the number

```
> 5+4
[1] 9
> 5-4
[1] 1
> 5*4
[1] 20
> 5/4
[1] 1.25
> 5^2
[1] 25
> 5%%2
[1] 1
> |
```

Relational Operators

- Used for comparison between values

Operator	Description
<	Less than
>	Greater than
<=	Less than or equal to
>=	Greater than or equal to
==	Equal to
!=	Not equal to

```
> 2 == 3
[1] FALSE
> 2 < 3
[1] TRUE
> 3 > 4
[1] FALSE
```

Questions or Comments