

Date: _____
Course 7.

Data Analysis With R

Computer PROGRAMMING:

Refers to giving instructions to a computer to perform an action or set of actions

WHY USE R, OVER SHEETS + SQL

"The skills were closely tied to the features and capabilities available in SQL, Tables & SHEETS ... But what if you want to work in more custom ways?"

It gives you additional functionalities.

R is used because

- **Popularity:** Frequently used for ana.
- **Tools:** Convenient library of ready-to-use tools for cleaning & analysis.
- **Focus:** Was created with stats in mind.
- **Adaptability:** Adapts well for use in ML & DA.
- **Availability:** OS. (open source.)

— — —
Small Advices...

"Errors and Error Messages are the part of the process"

"Be Persistent & Dig Deeper"

Sheets & SQL gives you the high level control over the data

Programming - Highest level of Control.

Programming lets you create your own applications and build your own function from scratch.

Syntax:

Shows you how to arrange the words & symbols you enter — so they make sense to computer.

Clarify the steps of Anal.

Saves time

Reproduce & Share

Programming
allows

flex.

R

A programming language beg.
Used for stat. anal. and other
data anal.

R is based on another: S



1970 John ...

Ross Ihaka

Robert Gentleman

R is

Accessible:

- Good - Beginners

Data Centric:

- Specifically designed to make
Data easier - efficient - powerful

Open Source:



Community:



Some Commands

① Download:

`install.packages("name")`

② Load:

`library("name")`

③ View

`View(name)`

④ Summary

`summary(name)`

R - is truly important when you have lots of data spread across multiple tables.

Week 2

Basic Concepts in R:

- Functions • Comments • Variables • Data Types
- Vectors • Pipes

Function Arguments:

Information that a function in R needs in order to run.

Get instant help: `?func`

Comments: `#`

Vector

Is a data elements of the same type stored in a sequence in R.