

ANT 6973

DATA VISUALIZATION AND EXPLORATION

PLAN FOR TODAY

1. Introductions
2. Orientation
 - Syllabus and course website
 - R and RStudio
 - DataCamp
3. Complete your first in-class activity!

WHO AM I?

Dr. Fernando Campos

Department of Anthropology

Email: fernando.campos@utsa.edu

- Phone: 210-458-8635
- Office: MH 4.03.44

Office Hours:

- Tuesdays and Thursdays, 2:00 to 3:30
- By appointment
- Or take your chances and just drop in!



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.

Graduate school in Calgary, Canada.



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.

Graduate school in Calgary, Canada.



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.

Graduate school in Calgary, Canada.

Taught at Tulane University, New Orleans.



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.

Graduate school in Calgary, Canada.

Taught at Tulane University, New Orleans.

Postdoc at Duke University, NC.



WHO AM I?

Dr. Fernando Campos

Born and raised in San Antonio.

Graduate school in Calgary, Canada.

Taught at Tulane University, New Orleans.

Postdoc at Duke University, NC.

This is my first year at UTSA!



WHAT DO I DO?

Dr. Fernando Campos

I'm a biological anthropologist, studying the behavioral ecology of wild primates.





Nicaragua

Costa Rica

Panama



Nicaragua

Área de Conservación
Guanacaste

Costa Rica

Panama





WHITE-FACED CAPUCHIN



MANTLED HOWLER



BLACK-HANDED SPIDER MONKEY



Kenya

Amboseli
National Park

Tanzania

AMBOSELI BABOON RESEARCH PROJECT



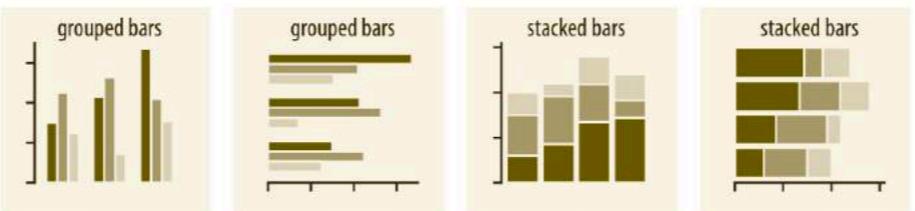
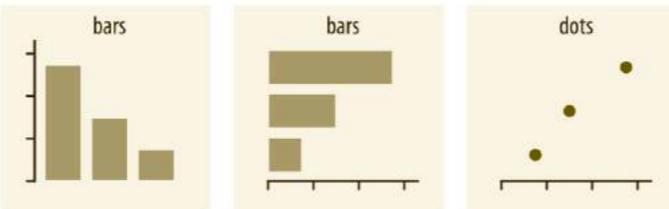




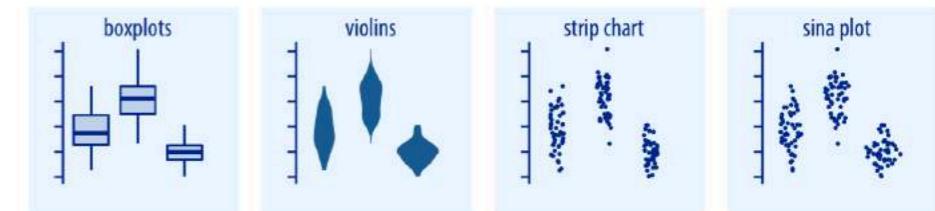
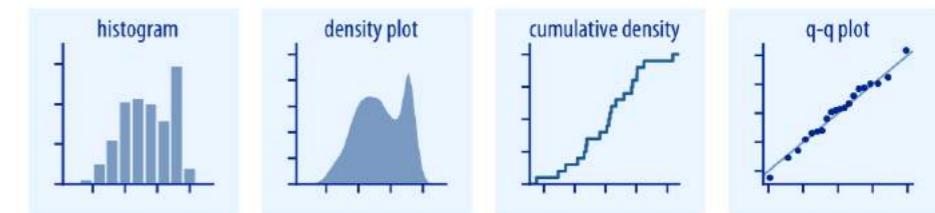
WHY AM I TEACHING
THIS COURSE?

DIRECTORY OF VISUALIZATIONS

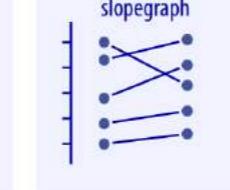
Amounts



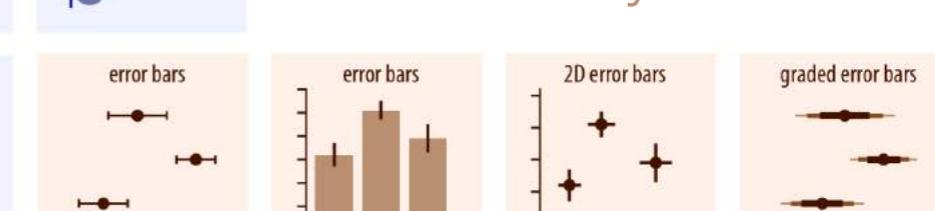
Proportions



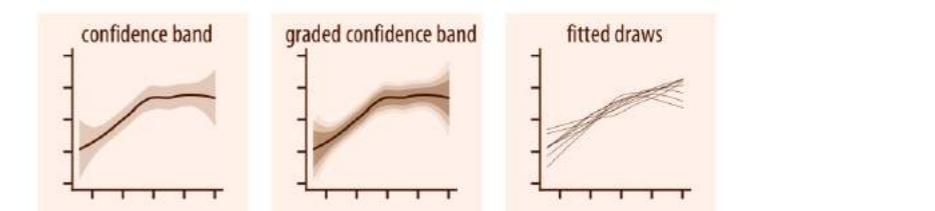
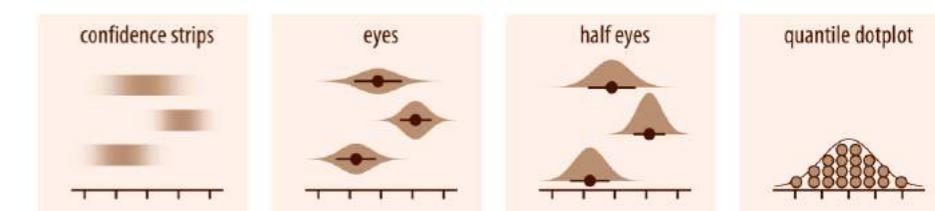
Distributions



Uncertainty



X-Y Relationships



YOUR TURN

1. Who are you?
2. What are your research interests?
3. What is your experience with data?
4. Do you have any experience using R, and if so, how much?
5. Is there anything specific that you hope to learn in this course?

**HELLO
my name is**

15:00

ORIENTATION

ORIENTATION

- Syllabus and course website
- R and RStudio
- DataCamp

COURSE WEBSITE

- <https://www.campos-lab.net/courses/ant6973-s2019/>
- Bookmark it!
- Go here for up-to-date information, schedule, assignments, readings, etc.
- I'll use Blackboard only for posting grades.
- Let's explore this a bit...

The screenshot shows a modern course website with a pink header. The header includes the course name "ANT 6937, Spring 2019" and navigation links for HOME, DATA CAMP, CHEAT SHEETS, SCHEDULE, and POLICIES. Below the header is a decorative background featuring abstract geometric shapes, a lightning bolt icon, and a scatter plot with a regression line. The main content area contains course details in a table:

Instructor	Dr. Fernando Campos ✉
Location	McKinney Humanities (MH) 3.02.52
Times	Thu 9:00 am - 11:45 am
Course	ANT 6937.002
CRN	34610
Office Hours	Tue/Thu, 2:00-3:30 pm, MH 4.03.44

Overview
This course is designed for graduate students in social or natural sciences who would like to gain experience in data visualization and analysis using R. It aims to be accessible to those who are completely new to programming, but also to provide new insights and skills to those who already have experience using R.

Approach
This course will use a mix of traditional and [flipped classroom](#) approaches, with new material introduced in reading assignments prior to class while class time will focus on applying those skills to data sets. The course will be activity-focused. Most concepts will be first introduced in assigned readings outside of class, leaving class time to focus on developing practical skills, completing activities, making progress on individual projects. We will cover topics in both data visualization as well as data manipulation. Students are encouraged to work collaboratively in and out of class. The [online course schedule](#) provides an overview of the topics covered as well as links to weekly readings, in-class activities, and any lecture material. This schedule is preliminary and always subject to change.

Texts
We will read sections from four openly licensed, online textbooks for this course. Hard copies of the books are *not* required, but some can be found at the links below.

- [Fundamentals of Data Visualization](#) by Claus Wilke.
- [Data Visualization: A practical introduction](#) by Kieran Healy. | [Hard copy](#).
- [ModernDive: An Introduction to Statistical and Data Sciences via R](#) by Chester Ismay and Albert Y. Kim.
- [R For Data Science](#) by Hadley Wickham and Garrett Grolemund. | [Hard copy](#).

Additional reading material will be linked from the [schedule](#).

At the bottom, there are links to GitHub, textbook resources (Wilke, Healy, ModernDive, R4DS), and a Creative Commons BY license logo.

ORIENTATION

- Syllabus and course website
- R and RStudio
- DataCamp

WHAT'S R VS. RSTUDIO?

- R is a programming language.
- RStudio is a nice user interface for R (an *integrated development environment* or IDE) that provides many convenient features and tools.



Engine



Dashboard

WHY R?



- R is the premier data analysis and visualization platform.

```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> █
```

WHY R?

- Pros:
 - Powerful, state-of-the-art
 - Used by many researchers—it is becoming a standard
 - Easy to extend, modify, and improve with add-on packages
 - Freely available for Unix, Windows & Mac
 - Programmable: if R can't do a particular task, you can program it.
 - Powerful plotting packages produce publication-quality graphics.



```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> |
```

WHY R?

- Cons:
 - Command-based rather than interactive in the point-and-click sense (like SPSS & Excel)
 - Significant learning hurdle
 - Specific commands easy to forget
 - Documentation sometimes poor
 - Still evolving: backward-compatibility has been an issue
 - Slow at times when compared to other programming languages



```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

The screenshot shows the CRAN homepage. At the top, there's a navigation bar with icons for back, forward, search, and other browser functions. The URL is https://cran.r-project.org. Below the bar, the title "The Comprehensive R Archive Network" is displayed. To the left, there's a large blue "R" logo with a grey outline. A sidebar on the left contains links for "CRAN Mirrors", "What's new?", "Task Views", "Search", "About R", "R Homepage", and "The R Journal". Another section lists "Software", "R Sources", "R Binaries", "Packages", and "Other". Documentation links include "Manuals", "FAQs", and "Contributed". On the right, there are several sections: "Download and Install R" (with links for Linux, Mac OS X, and Windows), "Source Code for all Platforms" (with a list of sources like releases, alpha/beta snapshots, daily patches, and older versions), "Questions About R" (with a note about license terms and frequently asked questions), "What are R and CRAN?" (with a brief description of R as a statistical language), "Submitting to CRAN" (with instructions for package submission), and a footer note about binary distribution compilation. The main content area has a light grey background with white text.

- Let's make sure we all have a current installation of R.
- <https://cran.r-project.org/>

- And a current installation of RStudio.
 - <https://www.rstudio.com/>
 - We'll get a crash course in how to use RStudio during the in-class activity.

g Download RStudio - RStudio X +

https://www.rstudio.com/products/rstudio/download/#download

R Studio Products Resources Pricing About Us Blogs

RStudio Desktop 1.1.463 — Release Notes

RStudio requires R 3.0.1+. If you don't already have R, download it here.

Linux users may need to import RStudio's public code-signing key prior to installation, depending on the operating system's security policy.

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab64d4
RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-10-29	a79032ba4d7daaa86a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	8a6755fa9fae2bafce289df3358aaf63
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-10-29	bc50d6bd34926c1cc3ae4a209d67d649
RStudio 1.1.463 - Ubuntu 16.04+/Debian 9+ (64-bit)	65 MB	2018-10-29	cf659db18619cc78d1592fefaa7c753
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-10-29	742f0bad60dfeaa3281576e14ad6699e
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-10-29	c7303067a0ca99deea7e427b856952d1

Zip/Tarballs

Zip/tar archives	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	122.9 MB	2018-10-29	1eb1d7758bd4bf4bb68d4a7c3fe8d894
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	90 MB	2018-10-29	ef9242b621d36c30de9d86b808840b41
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	98.3 MB	2018-10-29	5944b3dd118cfba46f2a6c484d768324
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.8 MB	2018-10-29	6bee446eeb4cc0e32967845e1f7ffe2b
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	91.4 MB	2018-10-29	3f63725769b7b976fdb754eb2a19108

Source Code

A tarball containing source code for RStudio v1.1.463 can be downloaded from here

250 Northern Ave, Boston, MA 02210
844-448-1212
info@rstudio.com

DMCA
Trademark
Privacy Policy
ECCN

 You can compile .R files just like .rmd files?!? The sneaky little notebook icon for doing this has been hiding... twitter.com/i/web/status/1... 12 hours ago

Try RStudio Server Pro for free!

Copyright 2018 RStudio | All Rights Reserved | Legal Terms



RStudio

Project: (None)

Console ~/ >

Environment History

Import Dataset List Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home Public

Name	Size	Mo
..		

A screenshot of the R Studio interface. The top bar shows the title "RStudio" and a "Project: (None)" dropdown. The main area has four panes: "Console" on the left, "Environment" and "History" tabs in the top right, "Files" and "Plots" tabs in the bottom right, and a file browser below them. The "Environment" tab shows "Global Environment" with a message "Environment is empty". The "Files" tab shows a directory structure with "Home" and "Public" folders. The "Plots" tab is currently selected, showing an empty plot area.

ORIENTATION

- Syllabus and course website
- R and RStudio
- DataCamp



DataCamp

- DataCamp is an online learning platform that focuses “data science,” which includes data visualization and data manipulation.
- It’s typically a paid service, but I have obtained licenses for you all.
- Let’s set it up.

The screenshot displays the DataCamp website with several key features highlighted:

- Top Navigation:** Includes links for "Learn", "Pricing", "For Business", and "Sign in".
- Hero Section:** Headline "THE SMARTEST WAY TO Learn Data Science Online". Subtext: "The skills people and businesses need to succeed are changing. No matter where you are in your career or what field you work in, you will need to understand the language of data. With DataCamp, you learn data science today and apply it tomorrow." Call-to-action button "Start Learning For Free".
- Logos:** Icons for Python, R, SQL, Spark, Git, Shell, and Spreadsheets.
- Account Creation:** Form titled "Create Your Free Account" with fields for LinkedIn, Facebook, Google, Email address, and Password, and a "Create Free Account" button.
- Mobile Learning:** Image showing a smartphone displaying a DataCamp exercise interface.
- Learning Anywhere:** Text: "DataCamp's lessons are bite-sized so you can learn in a way that fits your schedule, on any device. Tracks conveniently order the courses so you can find what fits your needs at a glance." Testimonial: "I think it is obvious that DataCamp is the best platform for those seeking to learn data science." - CAMERON WHITE USED DATACAMP TO HELP HIM LAND A JOB AS A DATA ANALYST.
- Code Editor:** Python code editor showing a script named "script.py" with code related to type conversion.
- Personalized Learning:** Text: "You learn best when you get to apply what you learn quickly. That's why we've developed Courses that combine short expert videos with immediate hands-on-the-keyboard exercises. Why we have built Practice Mode with instant and personalised feedback on every exercise. And why we have developed DataCamp Projects based on real-world problems."
- Bottom Navigation:** Icons for Python, R, SQL, Spark, Git, Shell, and Spreadsheets.



DataCamp

- Go to “DataCamp” tab on the course website.
- In the future, you can use the link to the ANT 6973 DataCamp Dashboard.
- For now, click the invite link.

ANT 6937 DataCamp Dashboard

Throughout the course we will be using DataCamp, a platform that provides an interactive way to learn how to use R to create visualizations and carry out data analysis.

Invite link: https://www.datacamp.com/groups/shared_links/860db90029f1891286789d3ece4f95d11df40484

[GITHUB](#) [TEXTBOOK: WILKE](#) [TEXTBOOK: HEALY](#) [TEXTBOOK: MODERNDIVE](#) [TEXTBOOK: R4DS](#)



DataCamp

- Create an account using your @utsa.edu email address.

The screenshot shows the DataCamp website's account creation page. At the top, there is a navigation bar with the DataCamp logo, a search bar, and links for "Sign in" and "Create Free Account". A message in the center of the bar says "Please register or log in first before joining the group". Below the bar, the main title "Create Your Account" is displayed. The form consists of five input fields: "First Name" (with a person icon), "Last Name" (with a person icon), "Email address" (with an envelope icon), "Password" (with a lock icon), and "Phone" (with a phone icon). A large yellow "Get Started" button is located at the bottom of the form. Below the form, a small note states: "By continuing you accept the Terms of Use and Privacy Policy, that your data will be stored outside of the EU, and that you are 16 years or older." To the right of the form, there is a link "Already have an account? Sign in". At the bottom of the page, there is a dark footer with sections for "LEARN", "RESOURCES", "PLANS", and "ABOUT". The "LEARN" section includes links for "Courses", "Skill Tracks", "Career Tracks", "Pricing", and "Become an Instructor". The "RESOURCES" section includes "Community", "RDocumentation", "Course Editor", "Upcoming Courses", and "Support". The "PLANS" section includes "For Business" and "For Classrooms". The "ABOUT" section includes "Stories", "Careers", "Contact", "Press", "Privacy Policy", and "Terms of Use". On the right side of the footer, there is a "DataCamp" logo and a brief description: "DataCamp offers interactive R, Python, Sheets, SQL and shell courses. All on topics in data science, statistics and machine learning. Learn from a team of expert teachers in the comfort of your browser with video lessons and fun coding challenges and projects." Below this, there is a "ABOUT THE COMPANY" section and social media icons for Facebook, Twitter, LinkedIn, and YouTube.

Create Your Account

First Name

Last Name

Email address

Password

Phone

Get Started

By continuing you accept the Terms of Use and Privacy Policy, that your data will be stored outside of the EU, and that you are 16 years or older.

Already have an account? [Sign in](#)

LEARN

Courses

Skill Tracks

Career Tracks

Pricing

Become an Instructor

RESOURCES

Community

RDocumentation

Course Editor

Upcoming Courses

Support

PLANS

For Business

For Classrooms

ABOUT

Stories

Careers

Contact

Press

Privacy Policy

Terms of Use

DataCamp

DataCamp offers interactive R, Python, Sheets, SQL and shell courses. All on topics in data science, statistics and machine learning. Learn from a team of expert teachers in the comfort of your browser with video lessons and fun coding challenges and projects.

ABOUT THE COMPANY

© 2019 DataCamp Inc

DATA CAMP

- Suggestion: turn off email notifications.

The screenshot shows the DataCamp homepage with a user profile for Fernando Campos (fernando.campos@utsa.edu) highlighted with a red box. The profile includes a recent activity log with two notifications: "New Course! Intermediate SQL Server" (2019-01-16 17:39:37 UTC) and "New Course! Human Resources Analytics in R" (2019-01-14 16:15:17 UTC). A green call-to-action box at the bottom left encourages users to answer questions to improve courses, with a "Get Started" button. The main content area features sections for "NEW COURSES" (Intermediate SQL Server, Hyperparameter Tuning in R, Human Resources Analytics in R: Predicting Employee Churn, Interactive Data Visualization with plotly in R) and "DAILY PRACTICE" (Joining Data in R with dplyr, Introduction to Python for Data Science, Introduction to R, Intermediate Python for Data Science). Below these are sections for "GET REAL LIFE EXPERIENCE WITH PROJECTS" (Introduction to DataCamp Projects, Exploring 67 years of LEGO, Dr. Semmelweis and the Discovery of Handwashing, Phyllotaxis: Draw flowers using mathematics), each labeled "NOT YET STARTED". A progress bar at the top right shows 0 / 250 XP Gained, 0 Day streak, and 12 Hours Left. A "Explore Projects Library" link is also visible.

What would you like to learn today?

Hi Fernando! You're enrolled in the **Data Visualization** track.

LEARN

Data Visualization with ggplot2 (Part 1)

Introduction 5 hours to go

PRACTICE

Joining Data in R with `dplyr`

Help us improve our courses by answering a few questions

Get Started

NEW COURSES

Intermediate SQL Server

Hyperparameter Tuning in R

Human Resources Analytics in R: Predicting Employee Churn

Interactive Data Visualization with `plotly` in R

Did you know DataCamp offers 213 courses? See more

DAILY PRACTICE

Joining Data in R with `dplyr`

Introduction to Python for Data Science

Introduction to R

Intermediate Python for Data Science

0 / 250 XP Gained | 0 Day streak | 12 Hours Left

Explore Projects Library

GET REAL LIFE EXPERIENCE WITH PROJECTS

Introduction to DataCamp Projects

Exploring 67 years of LEGO

Dr. Semmelweis and the Discovery of Handwashing

Phyllotaxis: Draw flowers using mathematics

NOT YET STARTED

NOT YET STARTED

NOT YET STARTED

NOT YET STARTED

https://www.datacamp.com/profile/account_settings

DATA CAMP

- Suggestion: turn off email notifications.

The screenshot shows the 'Account Settings' page on DataCamp. A pink box highlights the 'Notifications' section in the sidebar. Another pink box highlights the 'Don't send me any emails' button at the bottom of the page, which is turned on (green switch).

General
Password
Social
Notifications >
Payments
Advanced

Settings saved successfully

I want to receive emails about:

- Onboarding**: Tips and tricks to help you get started learning. (Switch off)
- Reminders**: Get notified when your subscription is set to renew. (Switch off)
- Course Launches**: Get notified when we launch new courses. (Switch off)
- Newsletters**: Receive DataCamp Weekly, our roundup of the best data science content on the web. (Switch off)
- Special Offers**: Be informed when there is a DataCamp promotion! (only a few emails per year) (Switch off)

Don't send me any emails
You only receive necessary emails like for a password reset. All the above will be turned off. (Switch on)

The screenshot continues from the previous one, showing other participation options. Both are turned on (green switches).

Help Improve DataCamp
I would like to participate in user research (interviews, surveys, user tests, etc.) to help improve DataCamp. This is done via email or banners within our applications. (Switch on)

Product Launches and Improvements
Inform me about the latest offers, product launches and improvements. This is usually done by showing you individualized ads on our website or on your social feeds. If you left your number, we might call or text you with a promotion or product update. (Switch on)

The footer navigation bar includes:

- LEARN: Courses, Skill Tracks, Career Tracks, Pricing, Become an Instructor
- RESOURCES: Community, RDocumentation, Course Editor, Upcoming Courses, Support
- PLANS: For Business, For Classrooms
- ABOUT: Stories, Careers, Contact, Press, Privacy Policy, Terms of Use

DataCamp logo: DataCamp offers interactive R, Python, Sheets, SQL and shell courses. All on topics in data science, statistics and machine learning. Learn from a team of expert teachers in the comfort of your browser with video lessons and fun coding challenges and projects.

ABOUT THE COMPANY

Download links: App Store, Google Play

Social media icons: Facebook, Twitter, LinkedIn, YouTube

© 2019 DataCamp Inc.



DataCamp

- Find links to DataCamp activities in course schedule under “Assignments”

ANT 6937, Spring 2019

ANT 6937, Spring 2019

HOME DATACAMP CHEAT SHEETS SCHEDULE POLICIES

⚠ PLEASE NOTE: Syllabus content will be updated frequently as we progress through the topics.

Date	Topics	Readings	Assignments	Slides	Projects
R, RSTUDIO, AND THE TIDYVERSE					
2019-01-17	Introduction & Setup				
2019-01-24	R and RStudio Basics				
DATA VISUALIZATION					
2019-01-31	Visualization Basics				
2019-02-07	Visualizing Amounts and Parts of a Whole				
2019-02-14	Visualizing Distributions and Uncertainty				
2019-02-21	Visualizing Relationships and Change Over Time				
2019-02-28	Principles of Figure Design 1: Coordinate Systems, Color Scales, and Facets				
2019-03-07	Principles of Figure Design 2: Annotations, Themes, and Common Pitfalls				
2019-03-14	No class: Spring Break ☀				
DATA MANIPULATION					
2019-03-21	Data Import and Principles of Tidy Data				
2019-03-28	No class: AAPA Conference 🌎				
2019-04-04	Data Manipulation Basics				
2019-04-11	Grouping and Summarizing				
2019-04-18	Relational Data and Joins				
FINAL PROJECTS					
2019-04-25	Work on Final Projects				
2019-05-02	Work on Final Projects				
2019-05-05	Final Projects Due by End of Day (23:59:59)				

GITHUB TEXTBOOK: WILKE TEXTBOOK: HEALY TEXTBOOK: MODERNDIVE TEXTBOOK: R4DS

CC BY

LET'S TAKE A SHORT BREAK

10:00