## CSDE 502 Winter 2022 Syllabus

#### **Instructor and Office Hours**

Phil Hurvitz
Raitt Hall 218 F
phurvitz@uw.edu
Office hours by appointment

#### **Time and Location**

Friday 10:30 AM to 12:20 PM

Zoom: https://washington.zoom.us/j/97609440755

## **Course Description**

This is a required course for students wishing to obtain a Demographic Methods Graduate Certificate from CSDE (<a href="https://csde.washington.edu/training/demographic-certificate/">https://csde.washington.edu/training/demographic-certificate/</a>). However, it is open to all interested students.

This course is meant to fill a perceived curriculum gap between methods courses that emphasize study design and statistics courses that teach statistical analysis. It focuses on applied methods for data preparation and will introduce the following topics: data management and documentation, data cleaning and variable creation, summarizing variables, working with demographic data, and reproducibility. In short, this course teaches introductory "data wrangling" focused primarily on demographic analysis applications.

The course is designed to be closely coupled with SOC/CSSS/CSDE 533 A (Research Methods in Demography). The techniques introduced in CSDE 502 will directly support the analytics taught in CSDE 533. Both courses will use R in the RStudio integrated development environment (IDE). In this course, particular emphasis will be placed on documentation to support transparency and reproducibility using the RMarkdown package.

## **Pre-requisites**

Introductory graduate level research methods and statistics, basic knowledge of R and RStudio. See <a href="https://csde.washington.edu/training/demographic-certificate/csde-502/#WinterQuarter">https://csde.washington.edu/training/demographic-certificate/csde-502/#WinterQuarter</a> for more details.

#### **Requirements and Grading**

This is a 2 credit, pass/fail course. In order to pass the course, students must satisfactorily complete the homework assignments and attend class. There will be weekly lectures and assignments to be completed outside of scheduled class times.

## **Dates**

January						February							March							
Su	Мо	Tu	we	Τh	Fr	Sa	Su	Мо	Tu	we	Τh	Fr	Sa	Su	Мо	Tu	we	Τh	Fr	Sa
						1			1	2	3	4	5			1	2	3	4	5
2	3	4	5	6	7	8	6	7	8	9	10	$1\overline{1}$	12	6	7	8	9	10	$1\overline{1}$	12
9	10	11	12	13	$1\overline{4}$	15	13	14	15	16	17	18	19	13	14	15	16	17	18	19
16	17	18	19	20	21	22	20	21	22	23	24	25	26	20	21	22	23	24	25	26
23	24	25	26	27	28	29	27	28						27	28	29	30	31		

# **Course Outline**

Quick introduction to RStudio, RMarkdown, file systems  Week 2: Jan 14  Introduction/review of basic R:  • data types • data structures • brief introduction to tidyverse • pipes (magrittr, tidyverse, native pipes)  Data • Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown: • code blocks • graphs • tables • equations • cross-references  Keyring (storing/accessing secrets such as passwords)  Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data  Metadata on data sets  Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18  Creating value labels  Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring	Course Outilile								
Week 2: Jan 14  Introduction/review of basic R:	Week 1: Jan 7	Quick introduction to CSDE computing							
data types     data structures     brief introduction to tidyverse     pipes (magrittr, tidyverse, native pipes) Data     Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown:     code blocks     graphs     tables     equations     cross-references     Keyring (storing/accessing secrets such as passwords) Data     Human Mortality Database     Human Fertility Database     Human Fertility Database     Human Fertility Database     Git: file versioning and code repository  Week 4: Jan 28  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data     Metadata on data sets     Ccmpp: Cohort Component Method of Population Projection Data:     Add Health public-use data  Week 7: Feb 18  Creating value labels     Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		Quick introduction to RStudio, RMarkdown, file systems							
• data structures • brief introduction to tidyverse • pipes (magrittr, tidyverse, native pipes) Data • Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown: • code blocks • graphs • tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18  Creating value labels Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring	Week 2: Jan 14	Introduction/review of basic R:							
brief introduction to tidyverse pipes (magrittr, tidyverse, native pipes)  Data Employee turnover data  Introduction/review of basic Rmarkdown: code blocks graphs tables equations equat		• data types							
• pipes (magrittr, tidyverse, native pipes) Data • Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown: • code blocks • graphs • tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18  Creating value labels Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring									
Data  • Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown:  • code blocks • graphs • tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18  Creating value labels Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		brief introduction to tidyverse							
Employee turnover data  Week 3: Jan 21  Introduction/review of basic Rmarkdown:     code blocks     graphs     tables     equations     cross-references     Keyring (storing/accessing secrets such as passwords)     Data     Human Mortality Database     Human Fertility Database     Human Fertility Database     Renvironments     R functions     Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data     Metadata on data sets     Ccmpp: Cohort Component Method of Population Projection     Data:     Add Health public-use data  Week 7: Feb 18  Creating value labels     Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		<ul> <li>pipes (magrittr, tidyverse, native pipes)</li> </ul>							
Week 3: Jan 21  Introduction/review of basic Rmarkdown:		Data							
• code blocks • graphs • tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Employee turnover data							
• graphs • tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring	Week 3: Jan 21	Introduction/review of basic Rmarkdown:							
• tables • equations • cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		• code blocks							
<ul> <li>equations</li> <li>cross-references</li> <li>Keyring (storing/accessing secrets such as passwords)</li> <li>Data</li> <li>Human Mortality Database</li> <li>Human Fertility Database</li> <li>Week 4: Jan 28</li> <li>R environments</li> <li>R functions</li> <li>Sampling in R</li> <li>Week 5: Feb 4</li> <li>Git: file versioning and code repository</li> <li>Week 6: Feb 11</li> <li>Reading labelled data</li> <li>Metadata on data sets</li> <li>Ccmpp: Cohort Component Method of Population Projection</li> <li>Data:         <ul> <li>Add Health public-use data</li> </ul> </li> <li>Week 7: Feb 18</li> <li>Creating value labels</li> <li>Tabulation (summarizing data)</li> <li>Week 8: Feb 25</li> <li>Scale scoring</li> </ul>		• graphs							
• cross-references Keyring (storing/accessing secrets such as passwords) Data • Human Mortality Database • Human Fertility Database • Human Fertility Database  Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data: • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		• tables							
Keyring (storing/accessing secrets such as passwords) Data      Human Mortality Database     Human Fertility Database  Week 4: Jan 28  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:     Add Health public-use data  Week 7: Feb 18  Creating value labels Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		-							
Data  Human Mortality Database Human Fertility Database  R environments R functions Sampling in R  Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  Add Health public-use data  Week 7: Feb 18  Creating value labels Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		• cross-references							
<ul> <li>Human Mortality Database</li> <li>Human Fertility Database</li> <li>Week 4: Jan 28         <ul> <li>R environments</li> <li>R functions</li> <li>Sampling in R</li> </ul> </li> <li>Week 5: Feb 4         <ul> <li>Git: file versioning and code repository</li> </ul> </li> <li>Week 6: Feb 11         <ul> <li>Reading labelled data</li> <li>Metadata on data sets</li> <li>Ccmpp: Cohort Component Method of Population Projection</li> <li>Data:</li></ul></li></ul>		Keyring (storing/accessing secrets such as passwords)							
● Human Fertility Database  Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  ● Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Data							
Week 4: Jan 28 R environments R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Human Mortality Database							
R functions Sampling in R  Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Human Fertility Database							
Week 5: Feb 4 Git: file versioning and code repository  Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring	Week 4: Jan 28	R environments							
Week 5: Feb 4  Git: file versioning and code repository  Week 6: Feb 11  Reading labelled data  Metadata on data sets  Ccmpp: Cohort Component Method of Population Projection  Data:  • Add Health public-use data  Week 7: Feb 18  Creating value labels  Tabulation (summarizing data)  Week 8: Feb 25  Scale scoring		R functions							
Week 6: Feb 11 Reading labelled data Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Sampling in R							
Metadata on data sets Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring	Week 5: Feb 4	Git: file versioning and code repository							
Ccmpp: Cohort Component Method of Population Projection Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring	Week 6: Feb 11	Reading labelled data							
Data:  • Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Metadata on data sets							
• Add Health public-use data  Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Ccmpp: Cohort Component Method of Population Projection							
Week 7: Feb 18 Creating value labels Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Data:							
Tabulation (summarizing data)  Week 8: Feb 25 Scale scoring		Add Health public-use data							
Week 8: Feb 25 Scale scoring	Week 7: Feb 18	Creating value labels							
		Tabulation (summarizing data)							
	Week 8: Feb 25	Scale scoring							
Reordering values		Reordering values							
Week 9: Mar 4 Miscellaneous data processing	Week 9: Mar 4	Miscellaneous data processing							
Week 10: Mar 11 Miscellaneous data processing, continued	Week 10: Mar 11	Miscellaneous data processing, continued							

## Class email list: csde502a\_wi22@uw.edu

Please send any questions you have about the course materials, assignments, programming, etc. to the class e-mail list. It is likely that more than one person will have the same question, therefore all students will be able to benefit from the question/answer/discussion. Questions sent directly to the instructor may be answered on the email list (with sender's identity removed). E-mail list archives are available at

https://mailman11.u.washington.edu/mailman/private/csde502a\_wi22/.

#### **Class materials** can be found in the course's Canvas site:

https://canvas.uw.edu/courses/1515226.

#### **Assignments**

Assignments are due the Thursday at 09:00 AM the week following when the assignment was distributed. *Late work will not be accepted without previous agreement by the instructor*. Each assignment will detail the file format to be submitted. Any assignment files to be turned in must be submitted by uploaded using Canvas, https://canvas.uw.edu/courses/1353992. Do not send any assignment materials to the instructor by e-mail.

## **Religious Accommodation**

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at Religious Accommodations Policy

(https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/). Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form (https://registrar.washington.edu/students/religious-accommodations-request/).

#### Conduct

The University of Washington Student Conduct Code (WAC 478-121) defines prohibited academic and behavioral conduct and describes how the University holds students accountable as they pursue their academic goals. Allegations of misconduct by students may be referred to the appropriate campus office for investigation and resolution. More information can be found online at <a href="https://www.washington.edu/studentconduct/">https://www.washington.edu/studentconduct/</a>.

## **Access and Accommodations**

Your experience in this class is important to me. If you have already established accommodations with Disability Resources for Students (DRS), please communicate your approved accommodations to me at your earliest convenience so we can discuss your needs in this course.

If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), you are welcome to contact DRS at 206-543-8924 or <a href="mailto:uwdrs@uw.edu">uwdrs@uw.edu</a> or <a href="mailto:disability.uw.edu">disability.uw.edu</a>. DRS offers resources and coordinates reasonable accommodations for students with disabilities and/or temporary health conditions. Reasonable accommodations are established through an interactive process between you, your instructor(s) and DRS. It is the policy and practice of the University of

Washington to create inclusive and accessible learning environments consistent with federal and state law.

#### **Academic Integrity**

The University takes academic integrity very seriously. Behaving with integrity is part of our responsibility to our shared learning community. If you are uncertain about if something is academic misconduct, ask me. I am willing to discuss questions you might have.

Acts of academic misconduct may include but are not limited to:

- Cheating (working collaboratively on quizzes/exams and discussion submissions, sharing answers and previewing quizzes/exams)
- Plagiarism (representing the work of others as your own without giving appropriate credit to the original author(s))
- Unauthorized collaboration (working with each other on assignments)

Concerns about these or other behaviors prohibited by the Student Conduct Code will be referred for investigation and adjudication by the office of the Dean of Arts and Sciences.

Students found to have engaged in academic misconduct may receive a zero on the assignment (or other possible outcome).

## **Safety**

Call SafeCampus at 206-685-7233 anytime – no matter where you work or study – to anonymously discuss safety and well-being concerns for yourself or others. SafeCampus's team of caring professionals will provide individualized support, while discussing short- and long-term solutions and connecting you with additional resources when requested.