



JOHNS HOPKINS

BLOOMBERG SCHOOL
of PUBLIC HEALTH

Survey Access & Analysis

Module 1
Lecture 3



Lecture Objectives

- ▶ Explain the main population-based surveys done routinely in the US and how to access the publically available data
- ▶ Understand basic survey sampling and design characteristics in terms of how they affect survey analysis
- ▶ Describe how to insert survey design and weighting information into Stata/R in preparation for analysis



JOHNS HOPKINS

BLOOMBERG SCHOOL
of PUBLIC HEALTH

Accessing Surveys

The material in this video is subject to the copyright of the owners of the material and is being provided for educational purposes under rules of fair use for registered students in this course only. No additional copies of the copyrighted work may be made or distributed.

Conceptual Framework for Chronic Diseases

There are ~120 CDC surveillance systems, many of which are geared towards chronic diseases and risk factors.

CDC'S NATIONAL CENTER FOR CHRONIC DISEASE PREVENTION AND HEALTH PROMOTION

Using Surveillance Systems to Prevent and Control Chronic Diseases



CDC surveillance systems collect data on chronic diseases and their risk factors. These systems—often the only source of such data—are vital for understanding how chronic diseases affect people in the United States. Without them, our prevention and control efforts would be guesswork. Surveillance data guide us in putting our resources to the best use.

Surveys

- ▶ Design
 - ▶ Data derived from national and State/provincial/territorial surveys
 - ▶ Cross-sectional or Longitudinal
- ▶ Information Received:
 - ▶ Risk factor data or biomarkers
 - ▶ Disease and disability (self report, medical record extraction, or physical exam)
 - ▶ Socio-economic data
 - ▶ Demographics

Surveys

► Advantages

- Population-based data for a range of health-related variables
 - Risk factors for most chronic diseases are shared
- Potential for case ascertainment
 - Self-reported diseases not requiring hospitalization
 - Retrospective data (before other surveillance was in place)

► Limitations

- Not linked to outcomes
- Certain self-reported data may have lower validity (*i.e.*, under-reporting of weight, or misunderstanding of a provider diagnosis)
- May systematically exclude key populations (hospitalized patients, incarcerated individuals, homeless individuals, rural areas, *etc.*)

Survey Examples

- ▶ Behavioral Risk Factor Surveillance System (BRFSS)
 - ▶ Telephone survey on risk factors and certain outcomes
- ▶ Youth Risk Behavior Surveillance System (YRBS)
 - ▶ Measures risk factors for leading causes of death, disability, and social problems among youth
- ▶ National Health and Nutrition Survey (NHANES)
 - ▶ Assess health and nutritional status of population
 - ▶ Combines interviews/physical exams/lab tests
- ▶ National Health Interview Survey (NHIS)
 - ▶ Large-scale household survey
- ▶ National Youth Tobacco Survey
 - ▶ School based measures of a range of tobacco products

Surveillance: Behavioral Risk Factor Surveillance System (BRFSS)

- ▶ Mostly assesses risk factors for chronic diseases by telephone survey
- ▶ Includes some infectious disease risk factors, such as HIV testing
- ▶ Started in 1984, now collects data all over US annually
 - ▶ In 2011, added cell phone sampling frame and began raking weights
- ▶ Designed at **national level**, but can get local (county)-level data
- ▶ Coordinated by the CDC, but done by each State Health Department
- ▶ Each state has the opportunity to add/remove certain questions
- ▶ SMART BRFSS:
 - ▶ BRFSS data is also analyzed at the level of Metropolitan/Micropolitan Statistical Area (MMSA) giving local health information
 - ▶ Can only give estimates if a certain response rate is achieved
 - ▶ Data available at about a one-year lag from the regular BRFSS data



Behavioral Risk Factor Surveillance System



Save the Date: BRFSS Annual Meeting April 23-28, 2023
View more information about the BRFSS annual spring meeting 2023



The Behavioral Risk Factor Surveillance System (BRFSS) is the nation's premier system of health-related telephone surveys that collect state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. Established in 1984 with 15 states, BRFSS now collects data in all 50 states as well as the District of Columbia and three U.S. territories. BRFSS completes more than 400,000 adult interviews each year, making it the largest continuously conducted health survey system in the world. [See More](#).

About BRFSS

BRFSS Questionnaires

Publications & Resources

Prevalence Data & Data Analysis Tools

Survey Data & Documentation

BRFSS Fact Sheets

Spotlight



[Working with Module Data?](#)
[BRFSS Statistical Briefs help guide research on many topics.](#)

Related Links

- [HealthData.gov](#)
- [Data Catalog](#)
- [CDC's Major Disease Surveillance Systems](#)



Looking for State-specific Information?
• [View State Coordinators](#)
• [State Data Uses](#)

The header features a search bar with a magnifying glass icon, navigation links for "Individuals and Families", "Providers and Professionals", "Media and Publications", and "Offices and Administrations". Below the navigation is the Maryland state flag and the text "MARYLAND DEPARTMENT OF HEALTH Surveys & Reports". On the right side of the header are social media icons for Facebook, Twitter, and Instagram, followed by an ellipsis.

Home Councils **Data** Contact Us

Surveys & Reports

About

Behavioral Risk Factor Surveillance System (BRFSS)

Data and Reports

Youth Risk Behavior Survey/Youth Tobacco Survey (YRBS/YTS)

About the Maryland YRBS/YTS

Behavioral Risk Factor Surveillance System

The Maryland Behavioral Risk Factor Surveillance System (BRFSS) is an ongoing telephone-based chronic disease surveillance program designed to collect data on the behaviors and conditions that place Maryland adults at risk for chronic diseases, injuries, and preventable infectious diseases. Maryland BRFSS also collects information on health care access and health disparities. The typical sample size is approximately 15,000 non-institutionalized Maryland residents age 18 and older per year.

The Maryland Department of Health, local health departments, other public health programs, universities, and research organizations use Maryland BRFSS data to assess need, plan programs, evaluate programs, inform policy, and apply for funding. Knowing the type and frequency of health issues and behavioral risk factors enables public health professionals to devise and implement programs geared toward the prevention of chronic diseases, injury, and disability, and to reduce health disparities.

Every summer, the Maryland Department of Health solicits requests for modules to be included in the next year's questionnaire. Organizations are able to request questions that are of interest for their programmatic needs. For more information on the module request process, please email the BRFSS and Data Manager at mdh.brfss@maryland.gov. Note: The request for modules for the 2023 questionnaire will be released in August of 2022.

The Maryland BRFSS is conducted under guidance from the Centers for Diseases Control and Prevention (CDC). For more information regarding the national BRFSS survey and its use in all 50 states and several territories, please visit www.cdc.gov/brfss/.

Maryland BRFSS data from 2011 through 2020 can be found on the [Maryland IBIS website](#) [under revision as of October 7, 2022]. Update 7.15.22 - The website is currently offline but will be back soon. Data can be analyzed by year, jurisdiction, and a number of demographic variables and other indicators. For all data requests not available through IBIS, please contact the BRFSS and Data Manager at mdh.brfss@maryland.gov.

Please consider participating in the BRFSS survey if you are called. The Maryland BRFSS program depends upon participation of Maryland residents of every age, sex, race, and ethnicity throughout the state in order to properly assess the health of Maryland residents. The survey is completely anonymous, and no identifying information will be recorded or shared.

If you have been contacted to participate in this survey and have questions or would like additional information, please contact the Maryland BRFSS program at 443-388-6740 or mdh.brfss@maryland.gov.

Additional Maryland BRFSS resources:

<https://health.maryland.gov/phpa/ccdpc/Reports/Pages/brfss.aspx>

Surveillance: Youth Risk Behavior Survey (YRBS)

- ▶ Focus on six risk behavior domains and some outcomes (asthma)
 - ▶ Contributions to injuries/violence
 - ▶ Alcohol/Drugs/Tobacco
 - ▶ Physical activity/Diet
- ▶ High-School based survey, started in 1990
 - ▶ There is a middle school version for certain states/cities
- ▶ Conducted **every other year** in the spring
- ▶ Includes publications on sexual minority youth, with relevant survey questions asked regularly at certain sites
- ▶ Generally national/state level, but data are available for certain large urban counties and certain tribal communities



Adolescent and School Health

CDC > DASH Home > Data & Statistics Home

Data & Statistics Home

Adolescent Behaviors and Experiences Survey (ABES) +

COVID Experiences Surveys (CovEx) +

National School COVID-19 Prevention Study (NSCPS)

Youth Risk Behavior Surveillance System (YRBSS) —

Overview

Results

Reports, Fact Sheets, and Publications

Participation

Frequently Asked Questions

Methods

Questionnaires

Data & Documentation

Trends Report

Toolkit

Journal Articles

Data Request Form

NYPANS

School Health Profiles +

Youth Risk Behavior Surveillance System (YRBSS)

[Print](#)

2021 YRBSS Data Release Coming Spring 2023



The Youth Risk Behavior Surveillance System (YRBSS) monitors six categories of health-related behaviors that contribute to the leading causes of death and disability among youth and adults, including—

- Behaviors that contribute to unintentional injuries and violence
- Sexual behaviors related to unintended pregnancy and sexually transmitted diseases, including HIV infection
- Alcohol and other drug use
- Tobacco use
- Unhealthy dietary behaviors
- Inadequate physical activity

YRBSS also measures the prevalence of obesity and asthma and other health-related behaviors plus sexual identity and sex of sexual contacts.

YRBSS is a system of surveys. It includes:

- a national school-based survey conducted by CDC and state, territorial, tribal,
- local surveys conducted by state, territorial, and local education and health agencies and tribal governments.

YRBSS Results



View 2019 YRBS national, state, and local school district results.

YRBSS Data



Access 2019 YRBS national, state and local school district data.

YRBS Explorer



Explore 2019 YRBS national, state, and local school district data via tables and graphs.

Trends Report



View trend data on health behaviors and experiences among U.S. high school students.

<https://www.cdc.gov/healthyyouth/data/yrbs/index.htm>

Surveillance: National Health Interview Survey (NHIS)

- ▶ In-person household interviews
 - ▶ Conducted through US Census Bureau investigators
- ▶ Started in 1957
- ▶ Topics include: Medical conditions, health insurance, clinical visits, health behaviors/risk factors
 - ▶ Consists of an annual core and a rotating core
- ▶ Traditionally has collected information from the entire household/family
 - ▶ Now changing to do a “sample” adult and child
 - ▶ Adding information on family relationships, family context, mental health information, etc.
- ▶ Data available through the National Center for Health Statistics

CDC NHIS - National Health Interview x +

cdc.gov/nchs/nhis/index.htm

 Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

A-Z Index

Search Search NCHS Advanced Search

National Center for Health Statistics

CDC > NCHS

 National Health Interview Survey

About NHIS

2019 Redesign +

What's New

Webinar

Questionnaires, Datasets, + and Related Documentation

Survey Reports and Data + Linked to NHIS

NHIS Early Release Program

Special Topics +

Native Hawaiian and Pacific Islander NHIS +

National Health Interview Survey

The National Health Interview Survey (NHIS) has monitored the health of the nation since 1957. NHIS data on a broad range of health topics are collected through personal household interviews. Survey results have been instrumental in providing data to track health status, health care access, and progress toward achieving national health objectives.

Survey Participants

If you were selected, learn more about participating

Survey Data and Documentation

Access data and documentation

Publications and Products

View NCHS publications using NHIS data

Interactive Data Query System

Generate estimates by select characteristics

Rediscover the National Health Interview Survey
Respected, Reliable, Redesigned

What's New

- Children Living in Households That

<https://www.cdc.gov/nchs/nhis/index.htm>

Population-Based Cohorts: NHANES

- ▶ National Health and Nutrition Examination Survey
 - ▶ Run by the CDC's National Center for Health Statistics
- ▶ Participants selected from the census, based on age/sex/race
 - ▶ Initiated in early 1960's, started as different groups/topics, but now more systematic
 - ▶ Has surveyed over 140,000 – about 5,000 per year with ~15 counties visited each year
 - ▶ Participants receive results of their physical exam and tests
- ▶ Combines in-person interviews and physical examinations (done in a mobile center) in adults and children (over-sample in ages 60+)
 - ▶ Includes blood/urine collection and storage, with periods of DNA collection
 - ▶ Data released for public use, and can enter into specific agreements for other data
- ▶ National Epidemiologic Followup Study (NHEFS): Follow up of the NHANES I cohort (1971-5) over several waves to follow for mortality, morbidity, and changes in risk factors.

NHANES - National Health and N +  CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

Search Search NCHS Advanced Search [A-Z Index](#)

 Paused

National Center for Health Statistics

CDC > NCHS

   

 National Health and Nutrition Examination Survey

Survey Participants

If you were selected, learn more about participating

Survey Data and Documentation

Access data, documentation & response rates

Publications and Products


Data Analysis Tutorials


National Health and Nutrition Examination Survey

Survey Participants

If you were selected, learn more about participating

Survey Data and Documentation

Access data, documentation & response rates

Publications and Products

Data Analysis Tutorials

<https://www.cdc.gov/nchs/nhanes/index.htm>



JOHNS HOPKINS

BLOOMBERG SCHOOL
of PUBLIC HEALTH

Survey Analysis

The material in this video is subject to the copyright of the owners of the material and is being provided for educational purposes under rules of fair use for registered students in this course only. No additional copies of the copyrighted work may be made or distributed.

Design Specifications: Simple vs Stratified Sampling

- ▶ Simple Random Sample:
 - ▶ Everyone in the population (of size n) has the SAME probability of being selected
 - ▶ Some exceptions to this, such as oversampling certain subpopulations
- ▶ Stratified Random Sample:
 - ▶ Partition your population into strata (groups such as by sex, age groups, etc.)
 - ▶ Then randomly sample from EACH stratum
 - Number in direct proportion to the size of the stratum: Allows all subgroups to contribute to your overall estimate
 - Number NOT in direct proportion to the stratum size, but instead oversample
 - ▶ Your populations are more homogenous within strata than across strata

Design Specifications: Cluster vs Stratified Sampling

- ▶ Cluster Sampling:
 - ▶ Partition your population into groups (these are your primary sampling units)
 - ▶ Usually done by geography
 - Often the goal is to reduce cost/effort – it's easier to control a group of data collectors in one district rather than across a large city
 - Note: Difficult to maintain heterogeneity of your populations
 - People within a cluster are often similar on several characteristics
 - ▶ You then sample the individuals within your randomly selected clusters
- ▶ Need to increase sample size to achieve the same precision
- ▶ Could introduce bias – if the majority of people in one cluster have a particular opinion, this gets generalized to the entire population

Design Specifications: Single vs Multi-Stage Sampling

- ▶ Single Stage:
 - ▶ Sample EVERYONE within your randomly selected clusters
- ▶ Multi-Stage:
 - ▶ Randomly select individuals from WITHIN each of your chosen clusters
- ▶ Stratified multi-stage sampling?
 - ▶ Stratify the country into urban/rural or another characteristic, then randomly select your clustering units (say, counties)
 - Clusters can be selected randomly, or by probability proportional to size
 - ▶ Then sample (say households) from within each cluster

Design Specifications

- ▶ Look on the survey website to obtain details on their design. Key things to look for:
 - ▶ Single or multi-stage sample?
 - If multi-stage, what was the Primary Sampling Unit (PSU)?
 - For NHIS, they divided the country in to a couple of thousand PSUs
 - ▶ Was there stratification?
 - For NHIS, they stratified the PSUs by size/certain SES indicators, and then sampled from those
 - ▶ What weights are you given? Design Weight? Post-stratification Weight?
 - ▶ Finite Population Correction (fpc): If you sample WITHOUT replacement from a small target (i.e., your sample is $>5\%$ of the target size), the distribution will no longer be normal. You include this, normally a number between 0 and 1. Gives better SE estimates
- ▶ Figure out what variable represents each of these from the codebook
- ▶ I often “cheat” and look at design specifications set by authors/lecturers to help me along

Setting Survey Specifications in your Program: NHIS Data

- ▶ For Stata: You set a “`svyset`” command, and from there forward Stata assumes this design/weights
- ▶ For R: You create a survey design object, which then gets specified in future commands
 - ▶ “nest” tells you if clusters are numbered uniquely within a stratum (i.e., start back at “1” in the next stratum)
- ▶ `svyset psu_p [pweight = wtfa], strata(strat_p)`
- ▶ `[design name] <- svydesign(ids=~psu_p, strat=~strat_p, weights=~wtfa, data=[insert name], nest=TRUE)`

Design Effect

- ▶ Quantifies how much the sample-to-sample variability of your survey is different from the expected variability of a hypothetical same-size dataset derived using simple random sampling
- ▶ Affects your variance (and therefore standard error) calculation
- ▶ Expressed as:
 - ▶ DEFF: ratio of designed-based variance over the hypothetical simple random sample variance
 - ▶ DEFT: ratio of standard errors
- ▶ Affects Chi-square calculation
 - ▶ Calculate the Rao-Scott Generalized Design Effect (Rao-Scott Adjusted Chi-Square)
 - ▶ Corrects the mean (1st order) and variance (2nd order) of the Pearson's Chi-square
 - ▶ Both Stata and R provide this (see coding supplement)

Survey Analysis

- ▶ See coding supplement for information on the following:
 - ▶ Estimate overall proportions, means, medians
 - ▶ Give accurate estimates for subpopulations
 - ▶ Create graphs of weighted data
 - ▶ Run regression models accounting for survey weights
- ▶ You will apply this information in the homework assignment
- ▶ Code is given for both Stata and R

Missing Data

- ▶ Not infrequently, you will have variables that are missing in your survey data
 - ▶ Intuitive – people may refuse to answer a question or provide a specimen
- ▶ This is referred to as “ITEM non-response”, differentiated from individuals who choose not to respond entirely (which we deal with using weighting)

Missing Data

- ▶ Survey weights are created using demographic information
 - ▶ Every individual has a weight
 - ▶ Your weights will sum to the population total – so each individual contributes a proportion towards the full population
- ▶ If you are missing a particular variable and drop those observations, the remaining weights will not add up to the population total
 - ▶ If you report a population prevalence using the “population total” that your statistical program reports, your prevalence will be wrong
 - ▶ This happens frequently in the literature!

Command: Stata: svydescribe (R: summary (design name))

```
. svydescribe sex
Survey: Describing stage 1 sampling units
```

pweight: wtfa_a
 VCE: linearized
 Single unit: missing
 Strata 1: pstrat
 SU 1: ppsu
 FPC 1: <zero>

Stratum	#Units included	#Units omitted	#Obs with complete data	#Obs with missing data	#Obs per included Unit		
			mean	max			
100	13	0	664	0	9	51.1	87
101	18	0	784	0	17	43.6	68
102	13	0	691	0	23	53.2	83
103	10	0	575	0	46	57.5	68
104	10	0	516	0	17	51.6	68
105	12	0	586	0	20	48.8	63
106	13	0	758	0	23	58.3	92
107	18	0	846	0	16	47.0	70
108	12	0	714	0	28	59.5	83
109	15	0	846	0	32	56.4	83
110	15	0	664	0	19	44.3	65
111	9	0	530	0	40	58.9	76
112	4	0	194	0	24	48.5	68
113	16	0	908	1	21	56.8	99
114	9	0	470	0	32	52.2	68
115	14	0	674	0	9	48.1	73
116	7	0	259	0	14	37.0	53
117	13	0	730	0	20	56.2	126
118	12	0	644	1	23	53.7	92
119	14	0	906	0	33	64.7	115
120	10	0	496	0	17	49.6	65
121	11	0	572	0	12	52.0	79
122	15	0	677	0	18	45.1	74
123	6	0	304	0	34	50.7	82
124	14	0	656	0	32	46.9	67
125	9	0	519	0	32	57.7	82
126	15	0	713	0	23	47.5	64
127	9	0	507	0	34	56.3	80
128	4	0	281	0	54	70.3	88
129	13	0	881	0	57	67.8	102
130	10	0	563	0	27	56.3	93
131	13	0	602	0	22	46.3	64
132	12	0	546	1	8	45.5	74
133	12	0	630	0	24	52.5	84
134	9	0	489	0	13	54.3	68
135	11	0	601	0	27	54.6	70
136	11	0	477	0	4	43.4	67
137	14	0	640	0	10	45.7	78
138	9	0	526	0	20	58.4	75
139	8	0	438	0	37	54.8	70
140	13	0	722	0	26	55.5	70
141	9	0	542	0	49	60.2	80
142	9	0	463	0	42	51.4	62
143	13	0	797	0	34	61.3	85
144	16	0	775	0	32	48.4	84
145	8	0	435	0	48	54.4	62
146	9	0	449	0	26	49.9	69
147	12	0	720	0	25	60.0	87
148	16	0	745	0	19	46.6	91
149	14	0	618	0	22	44.1	72
150	10	0	545	0	13	54.5	88
151	17	0	1,106	0	39	65.1	115
					31,994	3	
					31,997		

111	9	0	530	0	40	58.9	76
112	4	0	194	0	24	48.5	68
113	16	0	908	1	21	56.8	99
114	9	0	470	0	32	52.2	68
115	14	0	674	0	9	48.1	73
116	7	0	259	0	14	37.0	53
117	13	0	730	0	20	56.2	126
118	12	0	644	1	23	53.7	92

Command: svydescribe

Survey: Describing stage 1 sampling units							
		#Units included	#Units omitted	#Obs with complete data	#Obs with missing data	#Obs per included Unit	
Stratum						min	mean
100	13	0	655	9	8	50.4	86
101	18	0	763	16	17	42.4	68
102	13	0	676	15	22	52.0	82
103	10	0	556	19	42	55.6	67
104	10	0	505	11	17	50.5	66
105	12	0	570	16	18	47.5	61
106	13	0	746	12	23	57.4	90
107	18	0	832	14	16	46.2	68
108	12	0	701	13	28	58.4	82
109	15	0	835	11	32	55.7	83
110	15	0	646	18	19	43.1	65
111	9	0	505	25	40	56.1	72
112	4	0	192	2	23	48.0	67
113	16	0	884	25	21	55.3	98
114	9	0	458	12	32	50.9	67
115	14	0	648	26	9	45.3	69
116	7	0	458	12	32	50.9	67
117	13	0	648	26	9	46.3	69
118	11	0	715	15	20	56.1	72
119	12	0	633	12	22	52.8	98
120	10	0	480	16	17	48.0	63
121	11	0	561	11	12	51.0	79
122	15	0	665	12	17	44.3	73
123	6	0	298	6	6	49.7	82
124	14	0	638	18	31	45.6	65
125	9	0	518	1	32	57.6	82
126	15	0	697	16	22	46.5	63
127	9	0	499	8	33	55.4	80
128	4	0	280	4	1	50.0	88
129	13	0	850	31	54	65.4	99
130	10	0	537	26	26	53.7	91
131	13	0	589	13	22	45.3	64
132	12	0	536	11	8	44.7	71
133	12	0	622	8	24	51.8	84
134	9	0	478	11	11	53.1	67
135	11	0	590	11	27	53.6	70
136	11	0	467	10	4	42.5	66
137	14	0	620	20	10	44.3	76
138	9	0	513	13	20	57.0	75
139	8	0	436	2	36	54.5	69
140	13	0	694	28	26	53.4	67
141	9	0	521	21	48	57.9	80
142	9	0	443	20	41	49.2	58
143	13	0	778	19	33	59.8	85
144	16	0	759	16	29	47.4	80
145	8	0	423	12	46	52.9	61
146	9	0	442	7	26	49.1	68
147	12	0	709	11	25	59.1	86
148	16	0	725	20	19	45.3	91
149	14	0	610	8	21	43.6	72
150	10	0	537	18	13	52.7	84
151	17	0	1,068	38	36	62.8	112
		0	31,216	781	51.3	124	
			31,997				

Handling Missing Data

- ▶ How much missing data is “too much”?
 - ▶ Just as in study datasets, there is no absolute cut-off
 - ▶ How much do your estimates change?
- ▶ Possible techniques:
 - ▶ Impute missing values
 - ▶ Re-weight the complete cases so that they add up to the population total
- ▶ There is not consensus in the literature on how best to handle missing data
 - ▶ Langkamp *et al.* (Acad Pediatr, 2010) compared four commonly used methods and found multiple imputation and re-weighting to be the best of the options with >10% missing values.
- ▶ Both techniques often make reviewers “nervous”

Homework...Missing Data

- ▶ We will be “ignoring” missing data on the homework assignment
- ▶ Please enter commands as you would if the data were complete, and don’t worry about the slight differences in interpretation for population proportions or how you would have to deal with the missing data in order to conduct the other analyses

Next Session: Small Group Work

- ▶ Break off into groups of 4 – 6, based on chosen software (Stata vs. R)
- ▶ Work together, compare answers
- ▶ Please submit your final answers individually as a “quiz” in Courseplus
 - ▶ Versions for Stata and R are the same