LECTURE / SESSION[1] PLAN

Name Lecturer: Sofia Gil-Clavel	Date session: 5	Expected number of students: 20
Course title	R-Workshop	
Topic of lecture or session	Basic statistical analysis with R.	
Situational factors (e.g. group size, prior knowledge, expected motivation)	The class lasts 2hrs. The students are in different career stages, from PhD students to professors. The students have different backgrounds, from qualitative researchers that have never used R to quantitative researchers that want to learn new topics or move out of SPSS/STATA. The students use different Operating Systems (Windows, Mac, or Linux).	
Intended learning outcomes of this session	At the end of the session the students will be able to: 1. Understand the relevance of statistical modeling and situate it as part of the data science pipeline. 2. Translate statistical analyses to symbolic language. 3. Handle outliers 4. Fit basic linear models 5. Check basic assumptions and extract coefficients 6. Calculate confidence intervals 7. Use interaction terms 8. Perform analysis of variance 9. Fit prediction models 10. Situate inferential and predictive models in the universe of statistical models	

---

[1] A session means a teaching and learning session for university students in the bachelor or master. It can be a lecture, a seminar or a specific type of educational meeting with a group of students that is relevant for the discipline where the lecturer can demonstrate a whole range of teaching skills. It means that in one session each of the six didactic elements appears at least once. Your lesson plan should show that you are using a powerful learning environment and that students are activated.

Learning material (book, chapters, …)	The lecture is based on the book: • Silge, Max Kuhn and Julia. *Tidy Modeling with R*. Accessed February 7, 2025. https://www.tmwr.org/.
Media, equipment, tools	The students use their own laptops. The teacher needs access to a projector and a whiteboard.
Preparation for students	The students have access to the slides and codes before the class.

Time[2] (min.)	Didactic element (goal)[3] and topic[4]	What the teacher does[5] (teacher activity)	What students do (learner activity)	Evaluation[6] (feedback/assessment)
5	• Understand the relevance of statistical modeling and situate it as part of the data science pipeline.	• The teacher uses the slides to explain the relevance of statistical modeling and situate it as part of the data science pipeline.	• The students passively digest what the teacher is explaining.	• The students will use these concepts during the workshop. So, the teacher will detect when a student confuses them. Based on these confusions the teacher will be able to correct the student.
5	• Translate statistical analyses to	• The teacher uses the slides to explain why statistical models are considered symbolic language and	• The students passively digest what the teacher is explaining.	• The students will use these concepts during the workshop. So, the teacher will detect

---

	symbolic language.	explain that R was built to interpret and understand symbolic language.		when a student confuses them. Based on these confusions the teacher will be able to correct the student.
20	• Handle outliers	• The teacher uses the slides to introduce the data that will be used during the class. • The teacher opens RStudio and opens the already written script that will be used during the class. • The teacher uses the slides to explain the logic behind the code to remove the data outliers.	• The students passively digest what the teacher is explaining. • The students follow the steps in their computers. • The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
10	• Fit basic linear models	• The teacher uses the slides to explain the function lm and how it works together with a data frame and symbolic language.	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
20	• Check basic assumptions and extract coefficients	• The teacher uses the slides to refresh students memory regarding the linear models basic assumptions. • The teacher uses the slides to show the students how to extract	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a

		linear models coefficients and how to print statistical tests on their consoles.		common error, then the teacher uses the whiteboard to clarify.
10	• Calculate confidence intervals	• The teacher uses the slides to refresh students memory regarding what confidence intervals are. • The teacher explains the function confint.	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
10	• Use interaction terms	• The teacher uses the slides to refresh students memory regarding what an interaction term is. • The teacher explains how interaction terms translate into symbolic language and, therefore, into R.	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
10	• Perform analysis of variance	• The teacher uses the slides to refresh students memory regarding what analysis of variance is. • The teacher explains how to use the function anova.	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the

				teacher uses the whiteboard to clarify.
20	• Fit prediction models	• The teacher uses the slides to refresh students memory regarding what prediction models are. • The teacher explains how to use the function predict.	• The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides.	• The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
10	• Situate inferential and predictive models in the universe of statistical models	• The teacher uses the slides to situate inferential and predictive models in the universe of statistical models • Introduces an extra exercise that can be handed over any time the following week.	• The students passively digest what the teacher is explaining. • The floor is open for any final questions regarding the homework.	• The students will use these concepts to work on a homework that they can submit any time next week. • This homework is intended for those that need go deeper into the knowledge provided during this session.