

#### BBA182 Applied Statistics Week 1 (1)

DR SUSANNE HANSEN SARAL

EMAIL: <u>SUSANNE.SARAL@OKAN.EDU.TR</u>

HTTPS://PIAZZA.COM/CLASS/IXRJ5MMOX1U2T8?CID=4#

WWW.KHANACADEMY.COM





Basic course in statistical thinking and analysis. The primary goals are to help you:

Develop ability of statistical thinking and decision-making utilizing statistical tools in a context of business and management.

Acquire techniques to apply the proper current statistical tools to a broad range of business problems.

Topics covered include descriptive statistics and presentations, basic probability, various probability distributions, confidence intervals and hypothesis testing

Prerequisites: High school algebra



Students are expected to attend all scheduled classes as well as to bring all related course material in class (e.g. textbook, class notes, distribution tables, scientific calculator, etc.).

Students are liable to take the exams and participate in academic work (Khan Academy, Quiz and assigned homework) required for achieving the course.

Students who do not attend a minimum 70% of the classes (20 classes) will be considered as absent for the related course and therefore will get a VF



Students are permitted to arrive to the class in the first 15 minutes after the scheduled start of the course.

Students who arrive after 15 minutes of the scheduled start of the class will be considered absent.

Students who show up in the class after the break are considered absent.



#### ow I calculate your semester grade

Activities	<b>14 - week</b> s
Mid-term exam	30 %
Final exam	40 %
Participation on Khan Academy	25 %
and class quiz	
Class attendance	5%
Total	100%



#### Iculation of class attendance

Classes attended	Weight .10	Points
28 - 27	1.00	5
25 - 26	0.75	4
22 - 24	0.50	3
20 - 21	0.25	2
19 - 0	0.00	0 = VF



## Sharpe: Business Statistics, 3/e, Global Edition, Pearson

Newbold, Carlson, Thorne, Statistics for Business and Economics", 8<sup>th</sup> edition. (2012)



#### **Homework on Khan Academy**

Every week I will assign new homework on www.khanacademy.org

I give you a deadline and you will need to have mastered the homework in a weeks time.



Go to <u>www.khanacademy.org</u> create an account with your email address or your Facebook account (if you have one).

Add me (Susanne Hansen Saral) as a coach:

Follow the instructions from the hand-out



#### Piazza.com – class platform for:

## Posting class lectures, course syllabus, class announcement



#### Send me an email to the following address:

#### susanne.saral@okan.edu.tr



### What is statistics?

#### What is the average age of the students in this class-room?



### What is statistics?

Every statistical problem starts with a question!

- What was the overall customer satisfaction of Hilton Hotels in Turkey in 2015?
- How many pairs of jeans will GAP sell in the month of November
  2016 in Europe?
- How did you choose OKAN University for your studies?
- How many loafs of bread on average does a bakery store sell per

day?



Every statistical problem starts with a question!

Why would companies or individuals want to know the answers to these questions?



## To make good business decisions to help improve company revenues



### What is statistics?

#### How in Statistics do we go about answering such questions?

- O What was the overall customer satisfaction of Hilton Hotels in Turkey in 2015?
- How many pairs of jeans will GAP sell in the month of November 2016 in Europe?

• How did you choose OKAN University for your studies?



### We need to collect information from the source we are interested in to be able to answer such questions



### What is statistics?

#### Statistics concern **populations**

In the former examples the populations are :

All customers of Hilton hotels in Turkey in 2015

All pairs of jeans to be sold by GAP in Europe in November 2016

All students at OKAN University



### Statistical key definitions POPULATION

A **population** is the collection of all items of interest under investigation. **N** represents the population size

Populations are usually very large, therefore it is impossible to investigate entire populations. It would be too

- Time consuming
- Costly



### **Examples of Populations**

Incomes of **all families** in Izmir

All children in all elementary schools of a city

**All animals** in a farm

Human population on earth

**Total products** produced in one day in a factory



#### A sample is an observed subset of the population

• **n** represents the sample size

DR SUSANNE HANSEN SARAL



**ISTANBUL** 

### **Population vs. Sample**



#### Sample





### **Examples of Samples**

#### A **Sample** is a subset of the population

**A few parts,** of all parts produced selected, for testing defects

**10 children** from all elementary schools in a given city

The annual income of **33 families** out of all families in Izmir

The grade point average of **selected students** from OKAN University

**3 animals** out of a total of 25 animals



### Statistical key definitions PARAMETER VS. STATISTICS

A **parameter** is a specific characteristic of a population (mean, median, range, etc.)

Example: The mean (average) age of all students at OKAN

A **statistic** is a specific characteristic of a sample (sample mean, sample median, sample range, etc.)

Example: The **mean** (average) age of a sample of 500 students at OKAN



Populations are indefinite and their parameters are rarely known.

The only way we can find the estimated value of a population

parameter is by collecting a sample from the population of interest.



**Populations are usually infinite.** Therefore impossible to investigate the entire population

**Less time consuming** to investigate **a subset (sample)** of the population than investigating the entire population. Timely delivery of the results.

Less costly to administer, because workload is reduced

It is possible to obtain **statistical valid and reliable** results based on samples.



Our final objective in statistics is to make **valid and reliable** statements about the population in general based on sample data. (inferential statistics)

Therefore we need a sample that represents the entire population

One important principle that we must follow in the sample selection process is **randomness**.



#### Simple random sampling

Systematic sampling

Both techniques respect randomness and therefore provide **reliable** and **valid** data for statistical analysis



**Simple random sampling** is a procedure in which:

Each member/item in the population is chosen strictly by chance Each member/item in the population has an equal chance to be chosen Each member/item has to be independent from each other Every possible sample of n objects is equally likely to be chosen

The resulting sample is called a **random sample**.



In statistics we make decision about a population based on sample data, because the population parameter is unknown. Ex. Elections

Statisticians know that the sample statistic is rarely identical to the population parameter, but the two values are close.

The difference between the sample statistic and the population parameter is called **sampling error.** 



**Non-sampling errors**: Are errors not connected to the sampling procedure Population is not properly represented in the sample (Reader's Digest, 1936)

Survey subject may give incorrect or dishonest answer (because they did not understand the question or did not want to report the truth)

Survey subject fail to answer certain question in a survey (non response bias)

Subjects volonter to participate in a survey. Biased responses



# Drawing conclusion about a population based a sample information.



To draw conclusions about the population based on a sample we need to collect **data**.



#### Data = information

Data can be **numbers**: Size of a hotel bill, number of hotel guests, number of nights stayed in a Hilton hotel, size of a swimming-pool, etc.

Data can be **categories**: Gender, Nationalities, marital status, tourist attractions, codes, university major, etc.



Data are useless without a context.

When we deal with data we need to be able to answer at least the two following first questions in order to make sense of the data:

1) Who?

2) What?

2) When?

3) Where?

4) How?



#### Data values are useless without their context

Consider the following:

Amazon.com may collect the following data:

10675489	Ohio	10.99	Chris G.
Samuel P.	10783489	Katherine H.	Canada
16.99	Monique D.	11.99	15783947
15.98	Massachusetts	12837593	Illinois

What information can we get out of this?



We need to put the data into context in order to get information out of it

Purchase order #	Name	Price	Ship to State
10675489	Samuel P.	10.99	Ohio
10783489	Monique D.	16.99	Canada
15783947	Katherine H.	11.99	Massachusetts
12837593	Chris G.	15.98	Illinois



It is a basic study of transforming data into information :

how to collect it

how to organize it

how to summarize it, and finally

□ to analyze and interpret it



Market research

Survey (online questionnaires, paper questionnaires, etc.)

Interviews

□ Research experiments (medicine, psychology, economics)

Databases of companies, banks, insurance companies

Other sources



### Homework

Send me an email: <a href="mailto:susanne.saral@okan.edu.tr">susanne.saral@okan.edu.tr</a> TODAY

Create your Khan Academy account following the instruction of the hand-out

Go through the course syllabus

Watch the following YouTube video link: Introduction to Statistics

https://www.youtube.com/watch?v=BkV7D-fbKkQ