1. ETL process consists of Extract > Transform > Load. Which of these three processes is also known as data wrangling? **1 / 1 point**

- ◉ Transform
- ○ Load
- ○ Data wrangling is a term for another data warehouse process
- ○ Extraction

  ✓ **Correct**
  Correct, this process wrangles the data into the format suitable for destination and use.

2. The ELT process has no information loss. What is the main reason for this benefit? **1 / 1 point**

- ◉ Data replication
- ○ Separation between moving and processing data
- ○ Separates the data pipeline from processing
- ○ Data source integration

  ✓ **Correct**
  Correct, ELT provides a replica of the source data, and with that no information loss occurs.

3. ETL processes include a storage facility called a staging area. In ELT the staging area fits the description of what? **1 / 1 point**

- ○ Data warehouse
- ◉ Data lake
- ○ Electronic repository
- ○ Data mart

  ✓ **Correct**
  Correct, the staging area fits the description of a data lake, which is a modern self-serve repository for storing and manipulating raw data.

4. Which of the following pain points does ELT address? **1 / 1 point**

- ○ Cost effectiveness
- ◉ Challenges imposed by Big Data
- ○ Request for fixed processes
- ○ Lack of secure data

  ✓ **Correct**
  Correct, challenges like scalability imposed by Big Data are addressed.

5. There are many techniques for extracting data. Choosing the technique usually depends on what?

1 / 1 point

○ Type of client

○ Optical or analog

⦿ Intended use

○ Operating system

✓ **Correct**
Correct, it depends on the intended use of the data.

6. Extracting data from IoT devices involves large volumes of redundant data. What is used to decrease the data volume of redundant data and only extract features of interest from raw data?

1 / 1 point

○ APIs

○ Biometric sensors

⦿ Edge computing

○ SQL languages

✓ **Correct**
Correct, edge computing reduces the data volumes of redundant data by extracting features of interest from the raw data.

7. ETL uses the schema-on-write approach and ELT uses the schema-on-read approach. What is the biggest difference in these two approaches?

1 / 1 point

⦿ Limited versatility vs. versatility

○ Consistency

○ Stability

○ More data access

✓ **Correct**
Correct, the ETL approach limits the versatility, while the ELT approach is versatile since it obtains mulitple views of the same source data with ad-hoc schemas.

8. Which of the following examples of information loss during transformation can involve false negatives?

1 / 1 point

⦿ Edge computing

○ Filtering

○ Lossy data compression

○ Aggregation

✓ **Correct**
Correct, Edge computing devices, for example, false negatives appear in surveillance devices designed to only stream alarm signals, not the raw data.

9. Which of the following loading techniques is between batch and stream loading?

1 / 1 point

○ Incremental loading

○ On-demand loading

◉ Micro-batch loading

○ Parallel loading

✓ **Correct**
Correct, in between batch and stream loading, there is micro-batch loading.

10. Which of the following loading techniques can split a single file into smaller chunks?

1 / 1 point

○ Scheduled loading

◉ Parallel loading

○ Batch loading

○ Stream loading

✓ **Correct**
Correct, this technique splits single files into small chunks and loads them simultaneously.