1. Why should one use Azure Key Vault when working in the Azure environment? What are the pros and cons? What are the alternatives?
   a. Pros
      i. Security of storing sensitive information in credentials store which only the ADF service or Administrators can read from
      ii. If Credentials need to be rotated, ADF Linked Service will not need to be modified
      iii. When we migrate the ADF pipeline from Dev to Test to production, no change is necessary
   b. Cons
      i. Difficulty accommodating smaller company sizes
      ii. Fewer services compared to Amazon, HashiCorp, Google
   c. Alternatives
      i. AWS Certificate Manager
      ii. HashiCorp Vault
2. How do you achieve loop functionality within an Azure Data Factory pipeline? Why would you need to use this functionality in a data pipeline?
   a. Use a ForEach activity to iteratively execute functions over each unit within a larger directory e.g. files, folder, etc
3. What are expressions in Azure Data Factory? How are they helpful when designing a data pipeline? Please explain with an example.
   a. Expressions help filter parametrized values for ingestion without hardcoding
   b. Example: This expression checks if the 'name' attribute within the item has "old" or "testing" words in it, then uses "or" function to reverse it.
      `@substring(item().name,add(indexof(item().name,'_'),1),10)`
4. What are the pros and cons of parametrizing a dataset's activity in Azure Data Factory?
   a. Pros: more flexible ETL, dataset reusability in different pipelines or the same pipeline to access different files
   b. Cons: less readability for user the more complex the expressions are for parametrized values
5. What are the different supported file formats and compression codecs in Azure Data Factory? When will you use a Parquet file over an ORC file? Why would you choose an AVRO file format over a Parquet file format? (Reference)
   a. Supported file formats: Parquet, JSON, ORC, delimited text, binary, XML, Excel, AVRO
   b. Use Parquet file over ORC file when you have complex nested columns and only query a subset of the subcolumns often, and you plan to work with Spark.
   c. Use AVRO file over Parquet file when storing write-intensive data (row-based data format), you need high compatibility for old/new applications, you want to include the schema definition of your data as JSON text, you want to keep the data compact despite different schemas, and you plan to work with Kafka.