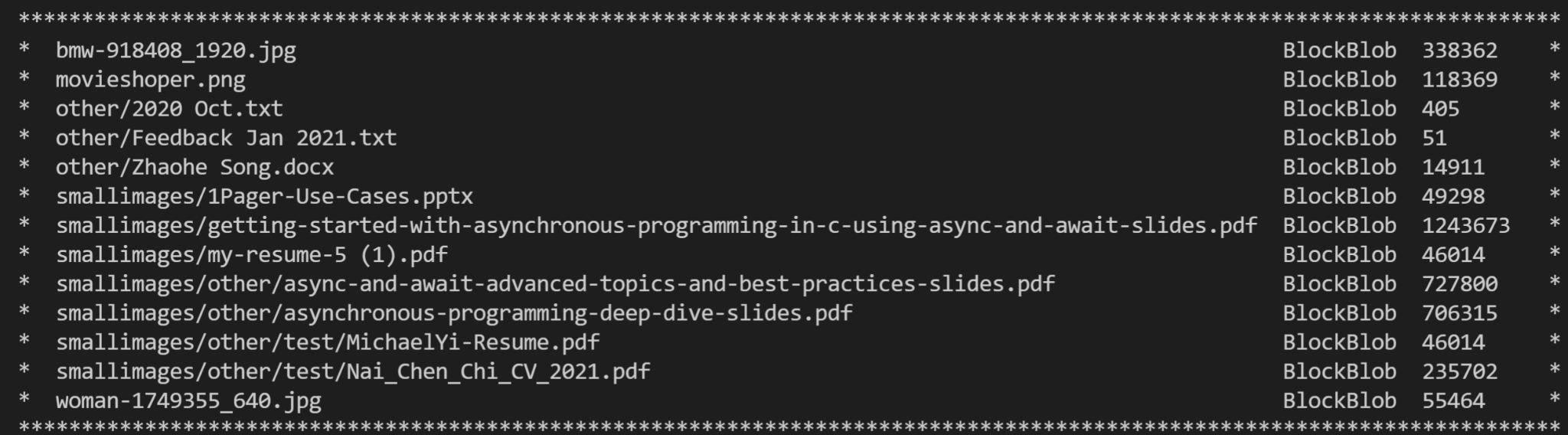
1. Create a script that will read and parse the given files and remove duplicates using python, then write back into a single CSV
   * When two rows are duplicates, they have the same information but might have different separators/casing. For example
     + “1234567890” instead of “123-456-7890”
     + “JANE” instead of “Jane”
     + “ Tom” instead of “Tom”
     + …
   * Once you clean up the anomalies, two rows that are supposed to be duplicates should have the exact same information/format.
2. Split movie.json into 8 smaller JSON files.
3. A paragraph on what PaaS, SaaS and IaaS are and the differences between them.
   * <https://www.bigcommerce.com/blog/saas-vs-paas-vs-iaas/#the-key-differences-between-on-premise-saas-paas-iaas>
   * <https://www.red74tech.com/advice/iaas-saas-paas-naas-different-flavors-of-cloud-services/>
   * https://apprenda.com/library/paas/iaas-paas-saas-explained-compared/
4. A paragraph on the differences between ETL and ELT. Also, list the pros and cons of each in a chart.
   * https://www.xplenty.com/blog/etl-vs-elt/
5. **(OPTIONAL)**Create a python script that will calculate/display:
   * Names, types and sizes of blobs in a certain container
   * Names and sizes of “folders” in a certain container
   * connection\_string = "DefaultEndpointsProtocol=https;AccountName=antrablobstorage;AccountKey=ECVP9sDWl64Ubd6w3lGd4d4fbiZuwHWWu1q/KoS2sCR18mwwkSxf1gLC7PvqCT1jWi3IYE87ZQtJYMIztIg3vg==;EndpointSuffix=core.windows.net"
   * container\_name = "imagescontainer"