2. S3 Select

I have uploaded the training set of the Titanic data set. By using the following `select s._2, s._5, s._6 from s3object s where s._2 = '1'` I have selected the columns Survived, Sex and Age and where the person has survived. The result was that the majority of the survived passengers where either female, children or of old age.
Alternatively, you can also use `select count(*) from s3object s` and count the number of rows in the csv file which in this particular case is 892.

4. S3 storage classes

There are six different Amazon S3 storage classes in total with each offering special designs for different use cases.
The first class is the S3 Standard which is for general-purpose storage and offers high durability, availability as well as performance object storage for frequently accessed data. Key features of the S3 standard are its low latency and high throughput performance, the durability of 99.999999% of objects across Availability Zones and its 99,99% availability over the year. You use this storage class if the access of data is known and is frequent.
The second class the Amazon S3 Intelligent-Tiering and is designed to optimize costs by automatically. It does so by moving data to the most cost-effective access tier without impacting performance. The storing of objects is done in two access tiers: one optimized for frequent access and another for infrequent access. This is one key feature of the storage class, where it moves objects between the two access tiers based on when it was last accessed. The features of the S3 Standard are also present in this storage class.
The S3 Standard-IA is for infrequent access and its speciallity is rapid access of data when needed. Then compared to S3 Standard the price per GB storage and per data retrieval. It is ideal for long term storage, backups and as storage for recovery files. The S3 One Zone-IA is basically a low budget version of the S3 Standard-IA and stores its data in one AZ as opposed to the minimum of three like in the standard S3 or S3 Standard-IA.
The last class of storage types is the S3 Glacier and the S3 Glacier Deep Archive which are used for archiving of data. S3 Glacier is a cheap alternative for data archiving and has different retrieval options ranging from a few minutes to hours. The S3 Glacier Deep Archive is aimed to be the lowest-cost storage class and should be used for long term storage of data that is not regularly accessed. The retrieval time is within 12 hours, which is really slow.

Bonus: CloudFront

HTTPS comes with additional charges when trying to host a static website from S3. This is not the case when using CloudFront as there are no additional charges for that. You can use CloudFront to serve your static website from S3. CloudFront is a web service offered by Amazon that has the function of speeding up the distribution of static and dynamic web content. User requests are forwarded to the edge locations that CloudFront uses to cut down on latency to deliver the content as fast as possible. CloudFront takes advantage of using the AWS backbone network and edge servers to improve the performance of the website.

The content of the website is stored using a S3 bucket like usual. One use case for using CloudFront is that live streaming of videos can be offered by the website and delivered in a very fast manner. It is smart to use CloudFront in cases where the userbase of the website is geographically distributed in order to take advantage of the edge locations all around the world efficiently. An additional advantage is that CloudFront takes and caches objects in the S3 bucket and sends them to the edge locations for even faster delivery.