

**The downloaded "2013 Docs and Data Sample Dataset" (data\_2013.zip) includes 266 CSV files which could be merged into one Flat File Source, and imported into one table of SQL server database (SQL Server Native Client 11.0). SQL Server Database Primary Data File (.mdf) size is 1.44 GB (1,551,892,480 bytes). From the data table properties, it contains the information: Data space as 1,024.625 MB; Row count as 5091767; and Data records as 5091501.**

**Based on the information from the BackBlaze website, a couple of tables created as below:**

- [drive days] for the number of drive days for each model**
- [failures] for has the number of failures for each model**
- [model count] for the number of drives for each model as of Dec 31 (Limit the table to drives 1.5TB or bigger, to exclude boot drives.)**

**Final table created as [annual failure rate] which shows the annual failure rate for each selected model (Only drives with a population of 45 or more as of Dec 31 are included). "Data Analyze Report Chart" and "Data Analyze Summary" tables are in next two pages.**

**The sample size is critically important for us to work with, and it helps us to understand certainty of results. It would make sense to filter few less count records from the final table and in conclusion, the divers with lower annual failure rate are:**

**Hitachi HDS5C3030ALA630**

**Hitachi HDS723030ALA640**

**Hitachi HDS722020ALA330**

**Hitachi HDS5C4040ALE630**

**The divers with higher annual failure rate are:**

**ST31500341AS**

**ST31500541AS**

**ST3000DM001**

**ST32000542AS**

**The diver with middle annual failure rate is ST4000DM000.**

**Prepared by Charlie (Qing) Lin**

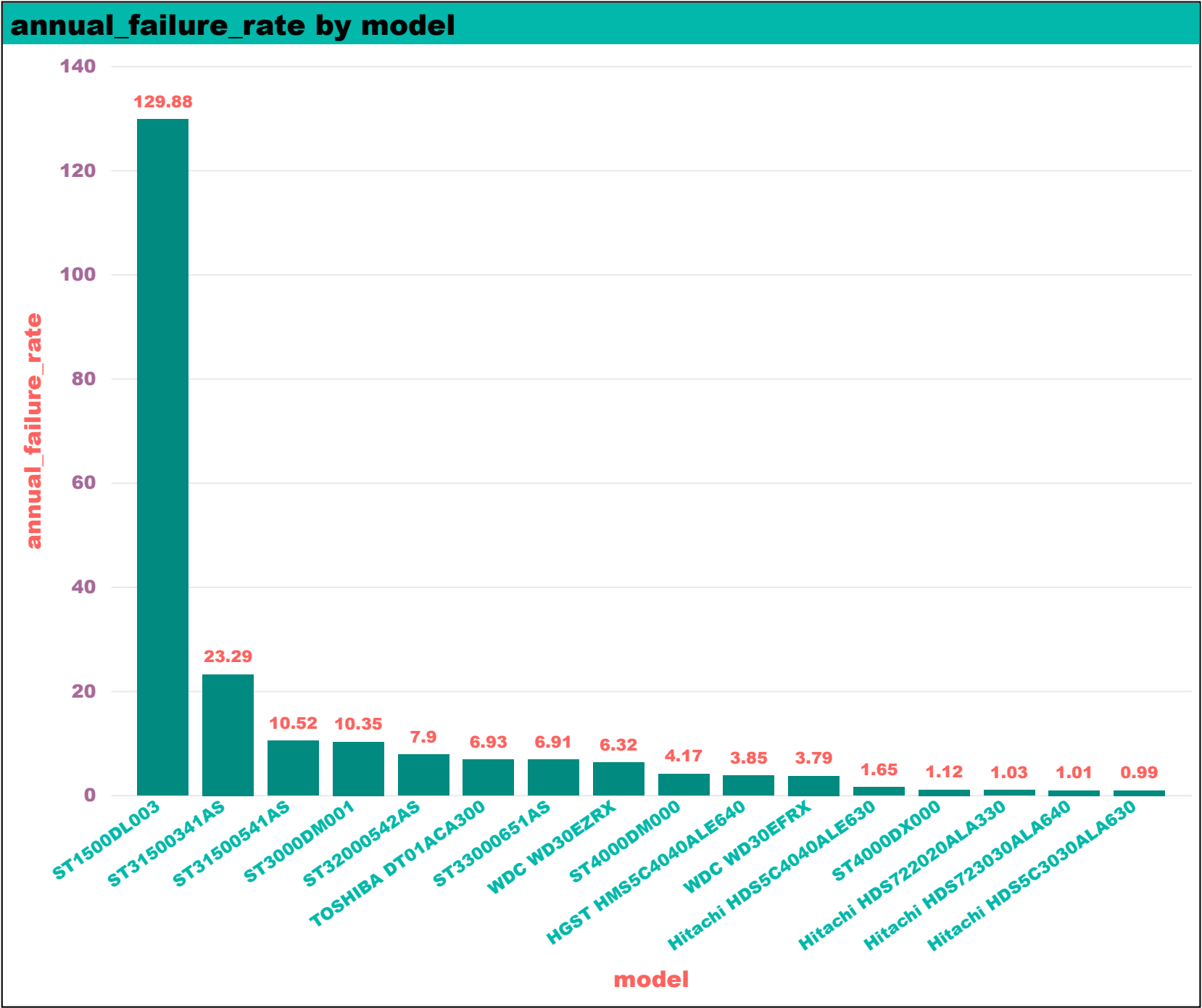
**Tel: 780-497-2805**

**email: qinglin@hotmail.com**

# Data Analyze Report

## Chart

model	annual_failure_rate
Hitachi HDS5C3030ALA630	0.99
Hitachi HDS723030ALA640	1.01
Hitachi HDS722020ALA330	1.03
ST4000DX000	1.12
Hitachi HDS5C4040ALE630	1.65
WDC WD30EFRX	3.79
HGST HMS5C4040ALE640	3.85
ST4000DM000	4.17
WDC WD30EZRX	6.32
ST33000651AS	6.91
TOSHIBA DT01ACA300	6.93
ST32000542AS	7.90
ST3000DM001	10.35
ST31500541AS	10.52
ST31500341AS	23.29
ST1500DL003	129.88



## Data Analyze Summary

model	brand	count	drive_sizes_GB	failures	annual_failure_rate ▲
Hitachi HDS5C3030ALA630	Hitachi	4592	3000	26	0.99
Hitachi HDS723030ALA640	Hitachi	1027	3000	6	1.01
Hitachi HDS722020ALA330	Hitachi	4716	2000	28	1.03
ST4000DX000	Seagate	179	4000	1	1.12
Hitachi HDS5C4040ALE630	Hitachi	2586	4000	23	1.65
WDC WD30EFRX	Western Digital	351	3000	4	3.79
HGST HMS5C4040ALE640	Hitachi	47	4000	1	3.85
ST4000DM000	Seagate	5200	4000	48	4.17
WDC WD30EZRX	Western Digital	388	3000	17	6.32
ST33000651AS	Seagate	293	3000	12	6.91
TOSHIBA DT01ACA300	Toshiba	58	3000	2	6.93
ST32000542AS	Seagate	288	2000	14	7.90
ST3000DM001	Seagate	4247	3000	266	10.35
ST31500541AS	Seagate	1929	1500	123	10.52
ST31500341AS	Seagate	539	1500	91	23.29
ST1500DL003	Seagate	51	1500	51	129.88