Data Analysis

Final Assignment Instructions

1. **Team Formation and Dataset**

Members: Rauch, Bilijesko, Frizberg

Group: Ohm\_Squad

Datasets: Westermo (processing for all, discussion: 2,5,7,9,10)

1. **Task Categories and Points**
   1. **Data Preprocessing and Basic Analysis**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | load-15m | sys-interrupt-rate | sys-thermal | disk-io-time | cpu-system | cpu-user | server-up | memory\_  used\_pct |
| count | 86401 | 86401 | 86401 | 86401 | 86401 | 86401 | 86401 | 86401 |
| mean | 0.1179 | 1135.7023 | 0.7040 | 0.0011 | 0.0505 | 0.0770 | 2.0000 | 3.4617 |
| std | 0.0869 | 725.0401 | 0.7168 | 0.0025 | 0.0384 | 0.0778 | 0.0000 | 0.4191 |
| min | 0.0000 | 653.6920 | 0.0000 | 0.0000 | 0.0140 | 0.0175 | 2.0000 | 3.0278 |
| 25% | 0.0500 | 1000.0000 | 0.0000 | 0.0008 | 0.0265 | 0.0340 | 2.0000 | 3.0816 |
| 50% | 0.1000 | 1048.5000 | 0.5000 | 0.0010 | 0.0370 | 0.0530 | 2.0000 | 3.1342 |
| 75% | 0.1700 | 1140.0500 | 1.3000 | 0.0014 | 0.0530 | 0.0765 | 2.0000 | 3.8060 |
| max | 0.7100 | 24490.3000 | 3.5500 | 0.3554 | 0.6755 | 1.9510 | 2.0000 | 5.3685 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | server-up | load-15m | Memory  \_used\_pct | cpu-user | cpu-system | sys-thermal | sys-interrupt-rate | disk-io-time |
| count | 71248 | 71248 | 71248 | 71248 | 71248 | 71248 | 71248 | 71248 |
| mean | 2.0000 | 0.1009 | 3.3662 | 0.0503 | 0.0368 | 0.6384 | 1046.9497 | 0.0009 |
| std | 0.0000 | 0.0762 | 0.3619 | 0.0225 | 0.0144 | 0.6831 | 80.1661 | 0.0005 |
| min | 2.0000 | 0.0000 | 3.0278 | 0.0175 | 0.0150 | 0.0000 | 874.5500 | 0.0000 |
| 25% | 2.0000 | 0.0400 | 3.0786 | 0.0325 | 0.0255 | 0.0000 | 992.0000 | 0.0006 |
| 50% | 2.0000 | 0.0900 | 3.0990 | 0.0440 | 0.0315 | 0.0500 | 1030.3500 | 0.0010 |
| 75% | 2.0000 | 0.1500 | 3.7712 | 0.0610 | 0.0460 | 1.2500 | 1082.8000 | 0.0012 |
| max | 2.0000 | 0.3500 | 4.8897 | 0.1400 | 0.0925 | 3.0506 | 1350.1000 | 0.0022 |

* 1. **Visualization and Exploratory Analysis**
  2. **Probability Analysis**
  3. **Statistical Theory Applications**
  4. **Regression Analysis**