If your data fits in memory there is no advantage to putting it in a database: it will only be slower and more frustrating.
"Databases are frustrating"

- Installation
- Maintenance
- Data Transfer

- Query Optimization
- Persistence
- Updates & Consistency
- Efficiency & Parallelism
In-Process
pip install duckdb

No Dependencies
Feature-Rich
Fast
<table>
<thead>
<tr>
<th>PID</th>
<th>USER</th>
<th>PRI</th>
<th>NI</th>
<th>VIRT</th>
<th>RES</th>
<th>S</th>
<th>CPU%</th>
<th>MEM%</th>
<th>TIME+</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>5676</td>
<td>hannes</td>
<td>25</td>
<td>0</td>
<td>392G</td>
<td>1418M</td>
<td>?</td>
<td>933.5</td>
<td>2.2</td>
<td>3:37.00</td>
<td>duckdb</td>
</tr>
<tr>
<td>80939</td>
<td>hannes</td>
<td>24</td>
<td>0</td>
<td>391G</td>
<td>139M</td>
<td>?</td>
<td>3.3</td>
<td>0.2</td>
<td>0:42.00</td>
<td>/System/Applications/Util</td>
</tr>
<tr>
<td>8872</td>
<td>hannes</td>
<td>17</td>
<td>0</td>
<td>390G</td>
<td>34512</td>
<td>?</td>
<td>2.1</td>
<td>0.1</td>
<td>0:00.00</td>
<td>/System/Library/CoreServices</td>
</tr>
<tr>
<td>8383</td>
<td>hannes</td>
<td>24</td>
<td>0</td>
<td>1523G</td>
<td>487M</td>
<td>?</td>
<td>2.0</td>
<td>0.7</td>
<td>0:20.00</td>
<td>/Applications/Discord.app</td>
</tr>
<tr>
<td>1853</td>
<td>hannes</td>
<td>17</td>
<td>0</td>
<td>389G</td>
<td>25504</td>
<td>?</td>
<td>0.9</td>
<td>0.0</td>
<td>5:18.00</td>
<td>/System/Library/CoreServices</td>
</tr>
<tr>
<td>8830</td>
<td>hannes</td>
<td>25</td>
<td>0</td>
<td>390G</td>
<td>12288</td>
<td>R</td>
<td>0.9</td>
<td>0.0</td>
<td>0:00.00</td>
<td>htop</td>
</tr>
<tr>
<td>8019</td>
<td>hannes</td>
<td>8</td>
<td>0</td>
<td>390G</td>
<td>48416</td>
<td>?</td>
<td>0.3</td>
<td>0.1</td>
<td>0:01.00</td>
<td>/Applications/Superhuman</td>
</tr>
<tr>
<td>1231</td>
<td>hannes</td>
<td>17</td>
<td>0</td>
<td>390G</td>
<td>31312</td>
<td>?</td>
<td>0.2</td>
<td>0.0</td>
<td>2:07.00</td>
<td>/System/Library/CoreServices</td>
</tr>
<tr>
<td>2068</td>
<td>hannes</td>
<td>17</td>
<td>0</td>
<td>391G</td>
<td>36672</td>
<td>?</td>
<td>0.2</td>
<td>0.1</td>
<td>1:30.00</td>
<td>/Applications/Rectangle.app</td>
</tr>
<tr>
<td>8021</td>
<td>hannes</td>
<td>8</td>
<td>0</td>
<td>1515G</td>
<td>276M</td>
<td>?</td>
<td>0.2</td>
<td>0.4</td>
<td>0:05.00</td>
<td>/Applications/Superhuman</td>
</tr>
</tbody>
</table>
Free
Demo

Code: https://tinyurl.com/ddbpdams1
One More Thing
“DuckSpark”
Demo
DuckDB

- An in-process analytical data management system
- Deeply integrated with Python, Pandas, Apache Arrow
- Questions?
- pip install duckdb

duckdb.org
discord.duckdb.org
@duckdb