# DeFi Liquidation Risk Analysis Report

## Overview

This report summarises a quick exploratory data analysis (EDA) and stress‐testing exercise on liquidations in the Aave v2 lending protocol. The goal is to understand how negative price movements in Ether (ETH) affect the frequency of liquidations and to provide quantitative estimates of liquidation risk under different shock scenarios.

The analysis draws on two main data sources:

**Liquidation events** – fetched from Etherscan via the Aave v2 pool contract. Logs are retrieved using the get\_latest\_block and chunked\_get\_logs functions from the src.api module. Approximately two million recent blocks are scanned for the LiquidationCall event.

**ETH prices** – daily OHLC data for the ETH/USDT pair is retrieved from Binance via the binance\_daily helper in src.prices. From this, daily returns are computed.

All data fetching, cleaning and merging logic has been moved into the src modules. The notebooks now simply call these functions, making them easier to read and maintain.

## Data processing

1. **Fetching the latest block** – The highest block number on Ethereum mainnet is obtained via get\_latest\_block(chainid=1, base=BASE, api\_key=API\_KEY). This avoids raw requests calls.
2. **Log retrieval** – Liquidation logs are pulled in chunks of 50 000 blocks using chunked\_get\_logs. Around two million blocks (~80–100 days of history) were scanned.
3. **Timestamp estimation** – Since logs provide block numbers but not precise timestamps, a cache of anchor blocks is built via the Etherscan getblockreward endpoint. A linear interpolation is then fitted to estimate timestamps for all blocks. This logic remains in the notebook for clarity.
4. **Daily aggregation** – The function liq\_daily\_count(df) groups the liquidation DataFrame by date, returning a time series of daily liquidation counts.
5. **Price data** – binance\_daily("ETHUSDT", 1000) fetches roughly 1000 days of daily price data. Daily returns are computed internally.
6. **Merging** – merge\_ret\_vs\_liq(daily\_px, daily\_liq) merges the price and liquidation series on the date column and drops missing values. The result is saved to data/eth\_ret\_vs\_liq.csv for reuse.

## Exploratory Data Analysis

The exploratory notebook (01\_quick\_eda.ipynb) visualises the cleaned data. Key plots produced include:

* **Daily liquidation counts** – a time series showing how frequently positions are liquidated on Aave v2. The series exhibits long periods of low activity punctuated by occasional spikes. See figs/daily\_liqs.png for details.
* **Scatter plots of ETH returns vs. liquidations** – several views (linear, jittered, and log‐scaled) highlight the relationship between the day’s ETH return and the number of liquidations. Days with large negative returns (≈−5 % or worse) correspond to dramatic increases in liquidations. Conversely, small positive or mildly negative returns tend to see very few liquidations. Examples are saved as figs/ret\_vs\_liq\_scatter.png and figs/ret\_vs\_liq\_jitter.png.
* **2D histograms** – heatmaps in both linear and logarithmic scales reveal the distribution of events across return/liquidation bins. They show that while most days cluster around modest returns and low liquidation counts, there is a long tail of high‐liquidation days triggered by sharp price drops. See figs/ret\_vs\_liq\_hist2d\_nolog.png and figs/ret\_vs\_liq\_hist2d\_balanced.png.

Quantitative summaries were also produced. For example, splitting days into “up” (ret ≥ 0) and “down” (ret < 0) categories reveals that more than 90 % of liquidations occur on down days. The average number of liquidations per day during price declines is an order of magnitude higher than during stable or rising markets.

## Stress testing methodology

To estimate liquidation risk under hypothetical market shocks, the second notebook (02\_stress\_test.ipynb) applies a kernel smoothing approach implemented in src.stress.kernel\_curve.

1. **Data preparation** – The merged return/liquidation DataFrame is sorted by date and cleaned. A 7‑day rolling sum of returns (ret7) is computed to capture multi‑day shocks.
2. **Shock grid** – A grid of price shocks from −30 % to −5 % (in 1 % increments) is defined. These values represent one‑day or multi‑day cumulative declines in ETH price.
3. **Kernel regression** – For each grid point, the kernel\_curve function estimates the expected liquidation count and constructs 25th and 75th percentile bands using bootstrap resampling. The smoothing bandwidth (sigma) controls how close in return space historical observations must be to influence the estimate.
4. **Results** – Two curves are produced:
5. **1‑day shock curve (curve1)** – uses daily returns (ret). A 5 % drop induces a modest rise in expected liquidations; larger shocks (>15 %) yield sharp increases.
6. **7‑day cumulative shock curve (curve7)** – uses the 7‑day rolling sum of returns. It is smoother because it averages over multi‑day movements, but it similarly shows that prolonged downturns greatly elevate liquidation risk.

Effective sample counts (the number of historical days with non‑negligible weight in each estimate) are plotted on a secondary axis. As shocks become more extreme, the number of comparable historical events drops, widening the confidence bands.

The combined plot of the two curves with confidence intervals and effective sample counts is saved as figs/stress\_curve\_compare.png. Individual stress‐curve CSV files (data/stress\_curve\_1d.csv and data/stress\_curve\_7d.csv) are also generated for further analysis.

## Key takeaways

Liquidation activity in Aave v2 is **highly sensitive to negative price shocks**. Days with returns worse than −5 % see a disproportionate number of liquidations.

The relationship between price shocks and liquidations is **non‑linear**. Small negative returns have little effect, but beyond a threshold (~−10 %) liquidation counts climb steeply.

Stress testing with kernel smoothing provides a **probabilistic estimate** of how many liquidations might occur under specified shock scenarios. This can inform protocol risk management and help users understand potential liquidation cascades during market stress.

## Deliverables

The repository now includes cleaned notebooks, modularised source code and analysis artifacts. To view or reproduce the analysis:

Open notebooks/01\_quick\_eda.ipynb for data extraction and EDA. It relies on helper functions in src/api.py, src/prices.py and src/features.py.

Open notebooks/02\_stress\_test.ipynb for the kernel regression stress test. It uses src/stress.py.

Figures (PNG files) are saved in the figs/ directory and CSV outputs are in data/.

This document (docs/report.md) provides a narrative overview of the workflow and key findings. Copy it into the deliverables/ folder or export it to PDF as needed for submission.