

*Module #3*

# Detecting abnormal markets: Early Warning Systems & C

**Raffaele Zenti** ([raffaele.zenti@virtualb.it](mailto:raffaele.zenti@virtualb.it))

Co-Founder, Head of Financial Data Science, Virtual B SpA

Market  
crashes do  
happen

### Largest Real Declines in U.S. Stock Market History

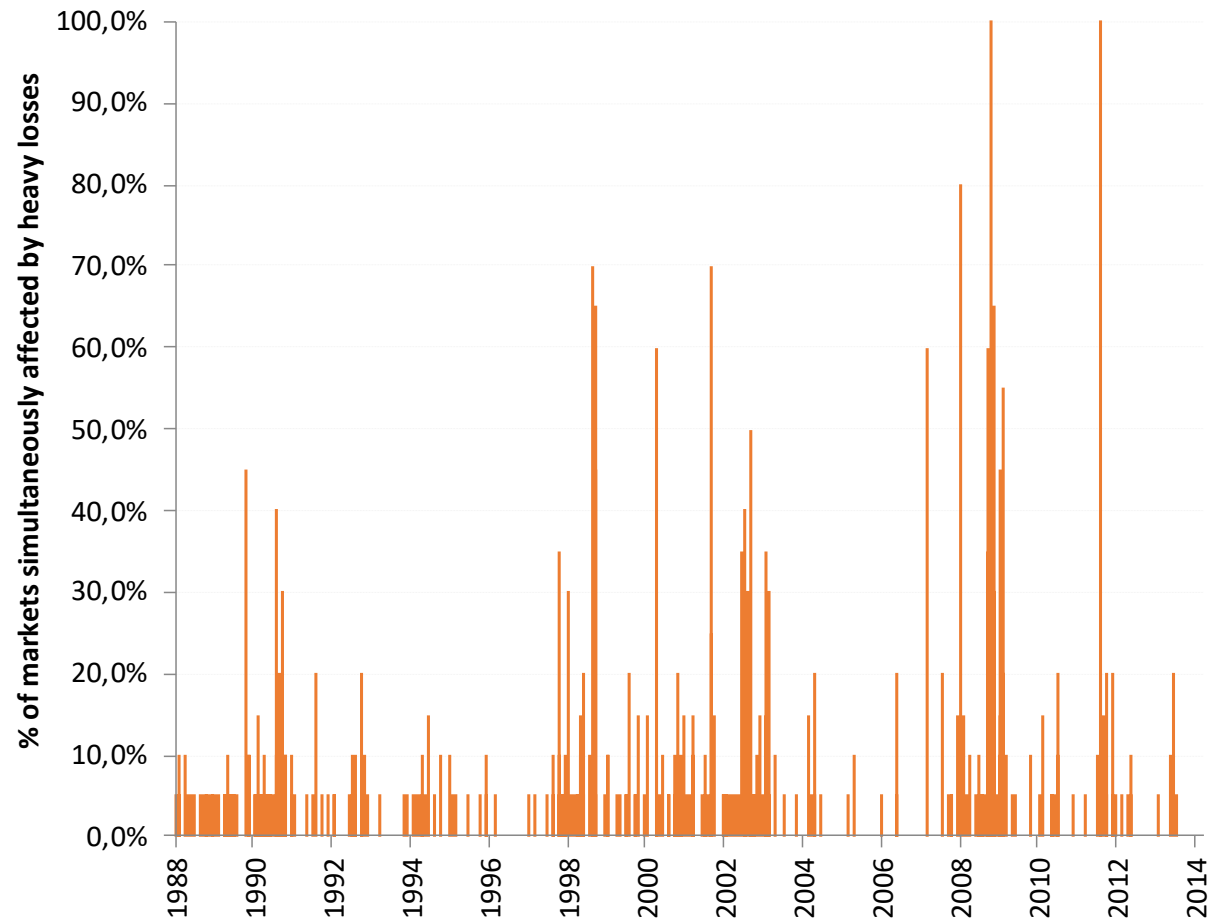
Pain Rank	Pain Index (%)	Peak	Trough	Recovery	Decline Rank	Decline (%)	Event(s)
1	100.00	Aug 1929	May 1932	Nov 1936	1	79.00	1929 Crash & Great Depression
2	89.34	Jun 1911	Dec 1920	Dec 1924	4	50.96	WWI & Influenza
3	85.51	Aug 2000	Feb 2009	May 2013	2	54.00	Lost Decade (Dot-Com Bust & Global Financial Crisis)
4	80.41	Dec 1972	Sep 1974	Jun 1983	3	51.87	Inflation, Vietnam, & Watergate
5	59.57	Feb 1937	Mar 1938	Feb 1945	5	49.93	Great Depression & WWII
6	29.06	May 1946	Feb 1948	Oct 1950	6	37.18	Postwar Bear Market
7	14.22	Nov 1968	Jun 1970	Nov 1972	7	35.54	Inflationary Bear Market
8	8.23	Jan 1906	Oct 1907	Aug 1908	8	34.22	Panic of 1907
9	8.18	Apr 1899	Jun 1900	Mar 1901	9	30.41	Cornering of Northern Pacific Stock
10	7.73	Aug 1987	Nov 1987	Jul 1989	10	30.21	Black Monday
11	6.25	Nov 1886	Mar 1888	May 1889	13	22.04	Depression & Railroad Strikes
12	5.00	Apr 1903	Sep 1903	Nov 1904	14	21.67	Rich Man's Panic
13	4.80	May 1890	Jul 1891	Feb 1892	17	20.11	Baring Brothers Crisis
14	3.55	Dec 1961	Jun 1962	Apr 1963	12	22.80	Height of Cold War & Cuban Missile Crisis
15	3.20	Aug 1897	Mar 1898	Aug 1898	15	21.13	Outbreak of Boer War
16	3.14	Oct 1892	Jul 1893	Mar 1894	11	27.32	Silver Agitation
17	3.11	Sep 1909	Jul 1910	Feb 1911	16	20.55	Enforcement of Sherman Antitrust Act
18	1.00	Dec 2019	Mar 2020	Jul 2020	18	20.00	COVID-19 Pandemic

Data as of Feb. 28, 2021. Sources: Kaplan et al. (2009); Ibbotson (2020); Morningstar Direct; Goetzmann, Ibbotson, and Peng (2000); Pierce (1982); [www.econ.yale.edu/~shiller/data.htm](http://www.econ.yale.edu/~shiller/data.htm), Ibbotson Associates SBBI US Large-Cap Stock Inflation Adjusted Total Return Extended Index.

Source: <https://www.morningstar.com/articles/1028407/in-long-history-of-market-crashes-coronavirus-crash-was-the-shortest>

The frequency of simultaneous "tail" events is increasing over time

- % of markets (all the 20 main world Stock Exchanges) simultaneously affected by large weekly losses (2.5%-tail of the empirical Copula estimated on weekly data since 1/1988)



Source: Zenti, R. (2014) «Volatility, decision models and complexity in financial market», Artificial Intelligence and Cognitive Science, Il Mulino

# The market's worst days had a big impact on investment returns

## The market's worst days have had a large effect on returns

Growth of \$1 in the S&P 500 Index from Dec. 31, 1927, to Dec. 31, 2015

Days	Ending value (\$)	Cumulative return (%)
Total cumulative return	115.40	11,440.46
Miss 10 best	38.28	3,728.33
Miss 10 worst	362.01	36,100.79
Miss 10 best and miss 10 worst	120.09	11,908.91
Cash	19.33	1,832.88

3X

Sources: Bloomberg L.P., Invesco, Morningstar.

# The problem

- One of the biggest problems of financial market is its annoying tendency to crash
- Market crises correspond to "risk-off" situations, in which risk premia and financial assets exhibit anomalous behavior
- There are big gains in detecting such crashes early on: risk prevention and improved financial performance
- Rather than predicting risk-off situations, it is sufficient to recognize them at their dawn
- It's more nowcasting than forecasting
- The large amount of financial data available invites us to solve the problem using data science

# Solution: Early Warning Systems

- Main goal: detecting crises before most damage has been made; and reducing false alarms
- Data Science provides us with many methods that can be successfully applied (also used in combination)



Let's look at  
the data

---



# Data overview

## Weekly data from Bloomberg

- Key equity indices
- Bond indices (Global, Corporate IG/HY, Inflation-linked, Municipals, Mortgages)
- Short/medium/long term interest rates
- Key exchange rates
- Commodities
- Leading indicators (Economic surprise, Baltic Dry Index)
- VIX (option implied volatility)

A label «abnormal/normal»






Brainstorming

*(don't be shy)*

- What would you do?
- What kind of algorithms would you use? Why?



# My focus: anomaly detection

---

- Anomaly: an observation which deviates so much from the other observations as to create suspicion that it was generated by a different mechanism
- Often indicative of something interesting
- Any deviations from the normal behavior that is unusual and significant is of special interest and may require action
- Detecting anomalies is increasingly becoming the core of many business operations, including finance

## Some applications

Early Warning Systems / Risk Modeling

Statistical Arbitrage (investments)

Transactional Fraud

Anti-Money Laundering

Customer Behavior Analytics (in general)



Matlab session starts |

# Take home on Early Warning Systems & Anomaly Detection

- Increasingly popular class of algorithms in Data Science
- You might use:
  - Probabilistic models (multivariate Gaussian, Copulas, whatever)
  - Classifiers
  - Clustering methods
  - Latent variable models like State-Space, PCA, etc
  - Simple proximity models, ie, distance-based model, like k-NN
  - ...
- You might also imagine an investment strategy based on that (buy/sell =  $f(\text{risk-on/off})$ )

## Now YOU

**It's your turn: use your favorite techniques to estimate customers' needs and recommend products, write the code (use my code, or start from scratch, or whatever...as you like), and we'll talk about it next time**

# Some useful links on Early Warning Systems for market crises

- <https://www.federalreservehistory.org/essays/stock-market-crash-of-1929>
- <https://www.federalreservehistory.org/essays/great-recession-and-its-aftermath>
- <https://arxiv.org/abs/0905.0220>
- [https://www.treasury.gov/initiatives/wsr/ofr/Documents/OFRwp0001\\_BisiasFloodLoValavanis\\_ASurveyOfSystemicRiskAnalytics.pdf](https://www.treasury.gov/initiatives/wsr/ofr/Documents/OFRwp0001_BisiasFloodLoValavanis_ASurveyOfSystemicRiskAnalytics.pdf).



## Next time

- Each group will present ideas, results, doubts, code snippet, etc
- Be short and concise (you are 20 groups):
  - 3' each group
  - Skip the details, go straight to the key methods and findings
  - Use selected plots, charts, tables (commented code snippets are good if concise)
  - If you have doubts and the minutes of the presentation are not enough, no problem, you can write to me (several of you have done so)
- Get ready to share your screen and your work - be it little or a lot – don't be shy
- My email address: [raffaele.zenti@virtualb.it](mailto:raffaele.zenti@virtualb.it)