

A better way to write regular updates using natural language processing

2021

Stephen Ryan

Market updates are essential

For clients, prospects, intermediaries and the media



Schroders

IN FOCUS

Quarterly markets review - Q2 2021

A look back at markets in Q2, when stocks made further gains and the US Federal Reserve hinted at future policy tightening.



Q2 market review: The road towards normalization

08 Jul 2021 | 10 min.



Investment Market Overview

Germany | 4th quarter 2020
Published in January 2021

2.2% ANNUAL DECLINE IN UK HOUSE PRICES IN DECEMBER	2.1% ANNUAL DECLINE IN PRIME CENTRAL LONDON PRICES IN JANUARY	0.8% ANNUAL RISE IN PRIME CENTRAL LONDON RENTS IN JANUARY	1.4% RISE IN UK RENTS IN YEAR TO DECEMBER
--	---	---	---



UK Residential Market Update

But what are updates made of?

Let's take a closer look at their contents



- ✓ Recurring updates usually comprise charts, images, **text** and possibly links
- ✓ **Text** is a mixture of statistical statements and commentary

Statistical statement



“The euro area seasonally-adjusted unemployment rate increased slightly to 7.8% in June, up from 7.7% in May...

Commentary



...which suggests that the social safety nets present in most European countries continue to prevent major job losses”

Sample of **text** from a market update

Drawn from monthly update by JPMorgan



Sample

Against this backdrop, new cases of Covid-19 have risen again in August but remain relatively low and the percentage of Covid-19 tests that are positive has remained well below the World Health Organization's recommended limit for reopening.

The Office for National Statistics reported a 20.4% (quarter on quarter) decline in GDP for the second quarter of the year. Consumption was strong in July with retail sales rising 3.6% over the month, rising back above their pre-crisis level.

However, with surveys showing that over 3 million workers are still furloughed, the unemployment rate could rise much higher with the scheme set to end in October and the government so far ruling out an extension.

These concerns about a fiscal cliff in October together with stalling Brexit negotiations, have weighed on investors' appetite for UK assets. Nevertheless, the FTSE All-Share still gained 2.4% in August, although it lagged the recovery in most other regions

Same text with statistical statements marked yellow



Sample

Against this backdrop, new cases of Covid-19 have risen again in August but remain relatively low and the percentage of Covid-19 tests that are positive has remained well below the World Health Organization's recommended limit for reopening.

The Office for National Statistics reported a 20.4% (quarter on quarter) decline in GDP for the second quarter of the year. Consumption was strong in July with retail sales rising 3.6% over the month, rising back above their pre-crisis level.

However, with surveys showing that over 3 million workers are still furloughed, the unemployment rate could rise much higher with the scheme set to end in October and the government so far ruling out an extension.

These concerns about a fiscal cliff in October together with stalling Brexit negotiations, have weighed on investors' appetite for UK assets. Nevertheless, the FTSE All-Share still gained 2.4% in August, although it lagged the recovery in most other regions

Statistical statements can make up half the total text

But producing them again and again is challenging



Sample

- ✓ Data comes from multiple sources (public and in-house)
- ✓ Process is laborious and prone to transcription error
- ✓ Opportunity cost: might time be better used?

Against this backdrop, new cases of Covid-19 have risen again in August but remain relatively low and the percentage of Covid-19 tests that are positive has remained well below the World Health Organization's recommended limit for reopening.

The Office for National Statistics reported a 20.4% (quarter on quarter) decline in GDP for the second quarter of the year. Consumption was strong in July with retail sales rising 3.6% over the month, rising back above their pre-crisis level.

However, with surveys showing that over 3 million workers are still furloughed, the unemployment rate could rise much higher with the scheme set to end in October and the government so far ruling out an extension.

These concerns about a fiscal cliff in October together with stalling Brexit negotiations, have weighed on investors' appetite for UK assets. Nevertheless, the FTSE All-Share still gained 2.4% in August, although it lagged the recovery in most other regions



The tyranny of the recurring update can lead to:

1. Bland and uninformative updates for clients and colleagues
2. Dull and repetitive work for researchers

There is a better approach

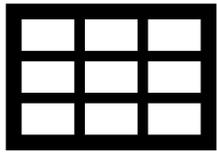
Computer generates the statistical text



1. Computer can read statistics and generate appropriate English
2. Latest available data is fed in via a spreadsheet
3. Appropriate English is generated instantly

Step by step description of the process

Data in spreadsheet → computer writer → human writer



Statistical data for the next regular update is saved in **spreadsheet**

Spreadsheet has just three columns:

- Column 1: Description of variable (such as “FTSE All-Share”)
- Column 2: Previous value of variable
- Column 3: Current value of variable



- **Computer** programme reads spreadsheet
- Compares the value in Column 3 to the value in Column 2
- Writes an appropriate statistical statement
- These statistical statements become the first draft of the regular update



- **Human** reads first draft of regular update
- Arrives at insightful interpretation of the statistical statements
- Converts insight into an appropriate commentary
- Statistical statements plus commentary become second draft

Conclusion

For regular updates, wrapping English around statistics is often half the work

- Regular updates are often dominated by charts and **text**
- Much of the **text** is a simple statistical statement
- It is time-consuming to produce and has an opportunity cost
- Natural language processing offers an alternative three-step approach

