# Guidelines from CERP for the Moodistan Portal

## Data Sources

For Moodistan, we propose the inclusion of two primary sources for mood analysis:

1. Newspapers and online publications
2. Social Media (Facebook, Twitter etc.)

These sources would constitute the basis for our mood formulation, updated on an hourly basis if possible or depending on feasibility, 4 hourly, 8 hourly, 12 hourly or daily.

## Natural Language Processing

The underlying technology would be developed by TPI to process natural language articles and social media posts in order to gauge the mood of Pakistanis within the context of different bins or themes/categories.

A proposed set of bins or themes is presented below:

1. Economy
2. Politics
3. Sports
4. Entertainment/Culture
5. Security/Conflict/Violence
6. Environment/Natural Conditions
7. Social

For each theme, we would present the mood in terms or intensity as a continuous spectrum from negative to positive and this ‘barometer’ should be visualized as a vertical barometer graphic.

## Analysis/Interface

### Infographic

In terms of the basic interface and capabilities, CERP envisions an embeddable infographic that presents daily or up to hourly values for each of the seven bins/themes defined above. So for each theme, we would see something like the pictures presented below. These are intended as a very rough guideline for design and is for TPI’s reference only. The barometers should be scaled in such a way that all seven bin/themes can be presented with their barometers within the same infographic (while it is preferable to have a vertical barometer, we can also go with horizontal barometers if these are easier to fit within a single infographic).

Figure-3 presents a rough layout for this infographic, we can see that all themes are represented and are clickable; the click behavior is that when a user clicks on a theme (let’s say ”Economy”), a dialogue bubble pops up with the top two or three articles/social media posts that were important in determining the related mood for the Economy theme. The user would then be able to open up the original articles using these links in order to read up more on these events.

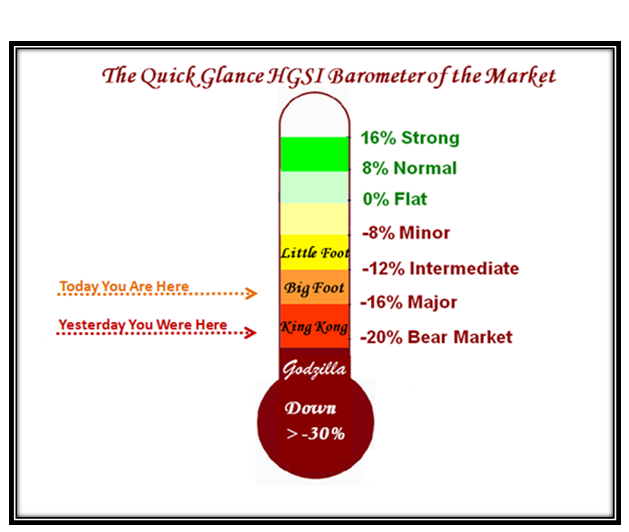


Figure-1

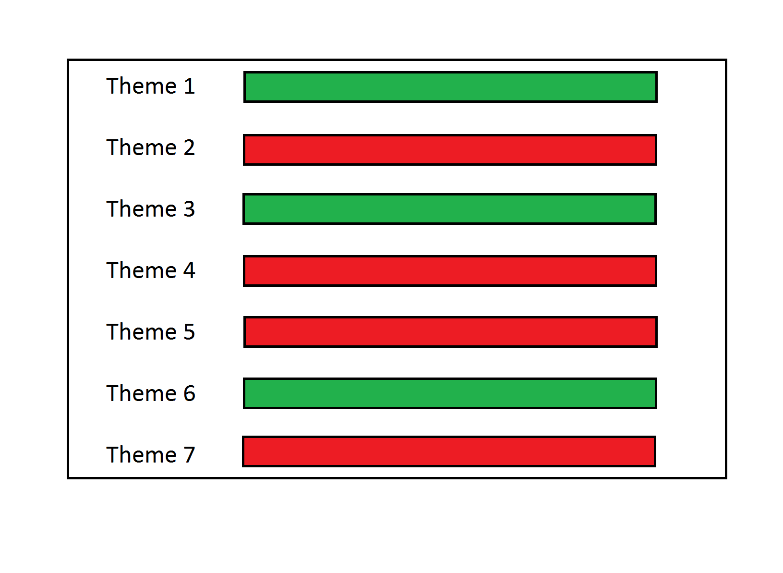


Figure-3

Colors are for reference only, in the actual infographic, we would have a continuous spectrum from negative (red) to positive (green) with an indicator line to show where along this spectrum the current mood lies.



Figure-2

### Trend Analysis

A second capability within this system is the ability to look at mood trends across time for one or more themes (out of seven) and the ability to add event lines to these trends to visualize the impact that a specific event had on the mood within the context of one or more themes.

The trend lines would also feature hover capabilities where the most relevant/important articles appear when the mouse hovers over a particular point on the trend line, similar to the infographic functionality described above. The user should also be able to input particular events and dates which would appear as vertical lines on the same graph as the trend lines. This would allow for a rudimentary form of event analysis wherein a user can see how a particular event affected the mood within the context of one or more themes.

A basic mockup of the trend analysis screen is presented below:

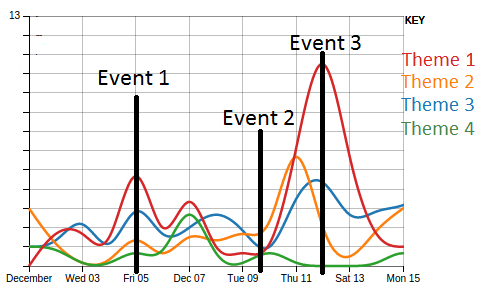


Figure-4

## Other Capabilities

There are some capabilities that CERP would like to see included in the portal, these can be divided into high and low priority capabilities. CERP would expect all high priority functionality to be developed, while low priority functions are optional depending on feasibility.

### High Priority

1. The infographic should be exportable as a PDF or JPEG with embedded CERP and TPI branding.
2. The infographic should be embeddable on other webpages.
3. The portal should include the ability to export raw data to users in csv or some other standard format.

### Low Priority

1. The portal may include the ability to allow users to upload their own articles or data sources for natural language processing in order to automatically recognize context and mood (based on our predefined moods) which the user is free to use as part of their own research or writing.