

Chapter 6: Storytelling Framework

Numbers have an important story to tell. They rely on you to give them a clear and convincing voice. –
Stephen Few

Learning Objectives

- Learn why storytelling is important.
- Understand the value of business storytelling.
- Understand what is data storytelling?
- Learn how to visualize narrative data and seven different data story types.
- Learn how to create analytics dashboards.

“If you want people to make the right decisions with data, you have to get in their head in a way they understand. Throughout human history, the way to do that has been with stories. [Stackpole, 2020]”

– Beth Stackpole on the book *The next chapter in analytics: data storytelling*

Importance of Business Storytelling

- Most communications lead to decision-making in business.
- Humans make their decisions based on both logic and emotion, but neuroscience has shown that decisions are primarily emotional, even in a boardroom.
- According to Portuguese American neuroscientist Antonio Damasio, emotions make better decisions [Damasio, 2009].

Importance of Business Storytelling

- Joshua Glenn and Rob Walker conducted one storytelling experiment known as **significant objects** [Glenn, 2009].
- They auctioned around 200 items purchased from a flea market on eBay.
- For the item description of those products, they asked different writers to write a short story for each item.
- The items they bought for \$1.25 per piece on average were sold for a total of \$8,000.
- This showed the power of storytelling; when a story resonates with the buyer's mind, it drives emotions which can transform insignificant objects into significant ones.

Storytelling Frameworks

1. Monomyth [Campbell, 2008] – This model of story structure, commonly known as the Hero's journey.

- It is extensively used in mythology and biblical stories throughout the history.
- There are three parts to monomyth model.
 1. An adventure.
 2. Challenges, which ultimately leads to a crisis.
 3. Finding a guide that helps him get through his crisis.

Storytelling Frameworks

2. Story Mountain [Anon, 2020] – This structure of the story is likely derived from Shakespeare's fine-art structure.

- A story mountain structure consists of five stages:
 1. Beginning,
 2. Conflict,
 3. Climax,
 4. Deflation, and
 5. Resolution.
- This structure is quite similar to the Hero's journey but unlike the Hero's journey, story mountain structure does not necessarily have a happy ending.

Storytelling Frameworks

- 3. Nested Loops** [Geljon, 2017] – In nested loops we tell stories within a story.
- Through different short stories, we describe the challenges, the findings, and the lessons learned.
 - This story structure is commonly used by television serial writers, who have a story to tell, but over a period of time through several episodes.
 - This structure of storytelling is useful when we do not have it all yet.
 - Hence, it will help in maintaining the audience's interest as they remain open to new things, throughout the story.

Data Storytelling

- The purpose of data storytelling is to communicate actionable insights from data through visual and narrative stories.
- Key elements of data storytelling?
 1. Skilfully connecting relevant data,
 2. Data visualizations, and
 3. Compelling narrative.
- A number of news organizations such as *The New York Times* or *Wall Street Journal* (WSJ) utilize smart and well- crafted data visualizations in their news articles to explain concepts.

Data Storytelling

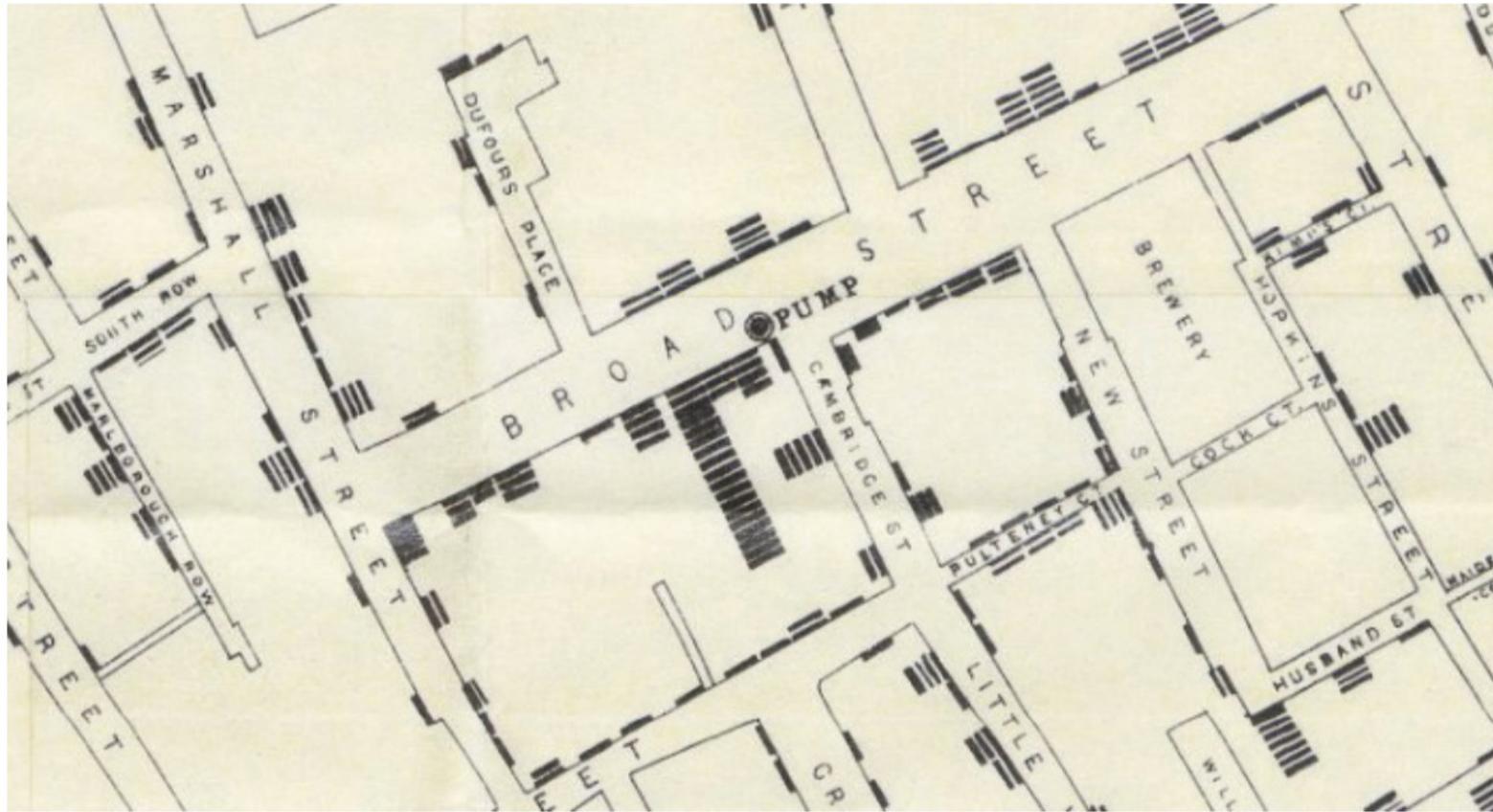


Figure 6.1 1854 cholera outbreak in London's Broad Street [DeBold, 2015].

Source: Battling Infectious Diseases in the 20th Century: The Impact of Vaccines. Available at <http://graphics.wsj.com/infectious-diseases-and-vaccines/>

Data Storytelling

- One such example is from 1854 (Refer to Fig. 6.1).
- Dr John Snow, a pioneer in disease mapping, visualized the 1854 cholera outbreak in London's Broad Street region [Snow, 1855].
- At that time, cholera was believed to be transmitted through the air.
- Snow mapped the cases on a city map.
- The map essentially represented each death as a bar, and there was a cluster of cholera cases around the Broadway Street pump.
- He used statistics to explain the connection between the number of cholera cases and the quality of the water source.

Narrative Types

Author-driven narrative

Author-driven Narrative

- Direction
- Linear Ordering
- Heavy messaging
- No Interactivity

- The method of presenting the findings along with static visualizations is known as “**Author-driven narrative**”.
- In this narrative, data and visualizations are chosen by the author and presented to the reader;
- The reader does not get to interact with the charts.
- Static charts used in presentation decks or printed magazine articles are examples of this kind of narrative.

Narrative Types

Reader-driven narratives

Reader-driven Narrative

- Discovery
- No prescribed Ordering
- No messaging
- Free Interactivity

- In a “**Reader-driven narratives**” we can curate data and provide it for free exploration by our readers.
- The reader can try out various innovative strategies to explore data and choose their own narratives.
- *Tableau public* cloud is one such example of readers building their own narrative on curated datasets.

- Additionally, there is a hybrid approach known as **narrative storytelling** [Segel and Heer 2010], which combines both communicative and exploratory visualizations.

Narrative Types

Narrative storytelling

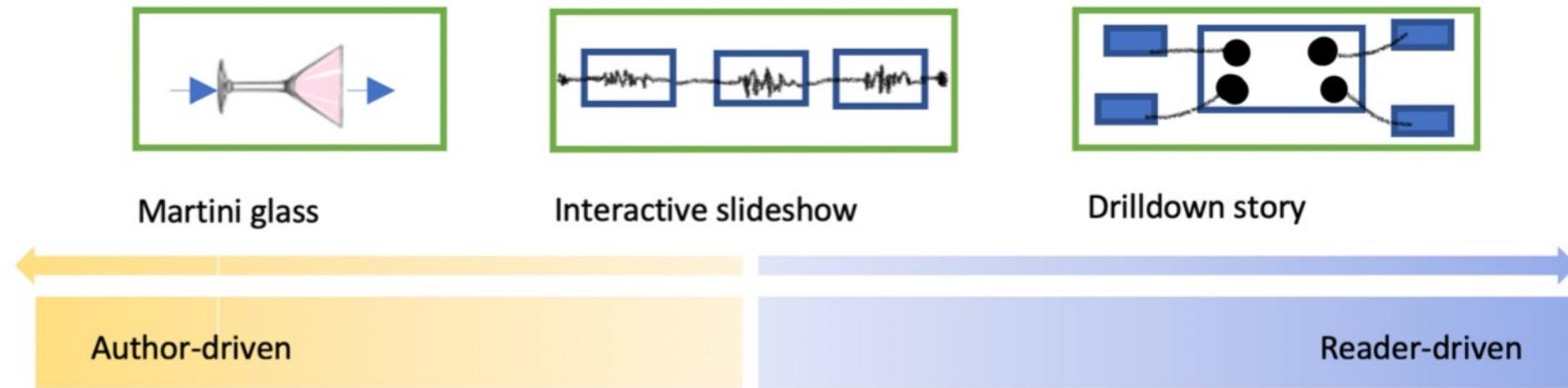


Figure 6.3 Narrative storytelling framework.

Source: Segel and Heer (2010)

- The authors introduce three narrative data storytelling frameworks:
 1. Martini-glass,
 2. Interactive slideshow, and
 3. Drilldown story

Narrative storytelling

- Narrative visualization normally consists of two approaches, one author-driven and one reader-driven approach, placed on two ends of a spectrum.
- Interactive visualizations, which can be built using varied tools like
 - Tableau,
 - Power BI,
 - Python/R visualization Libraries,
 - D3.js etc.
- Charts with interactive feature engages the audience to explore the data staying within the context of a more structured narrative.

Narrative storytelling

Published: February 2, 2010

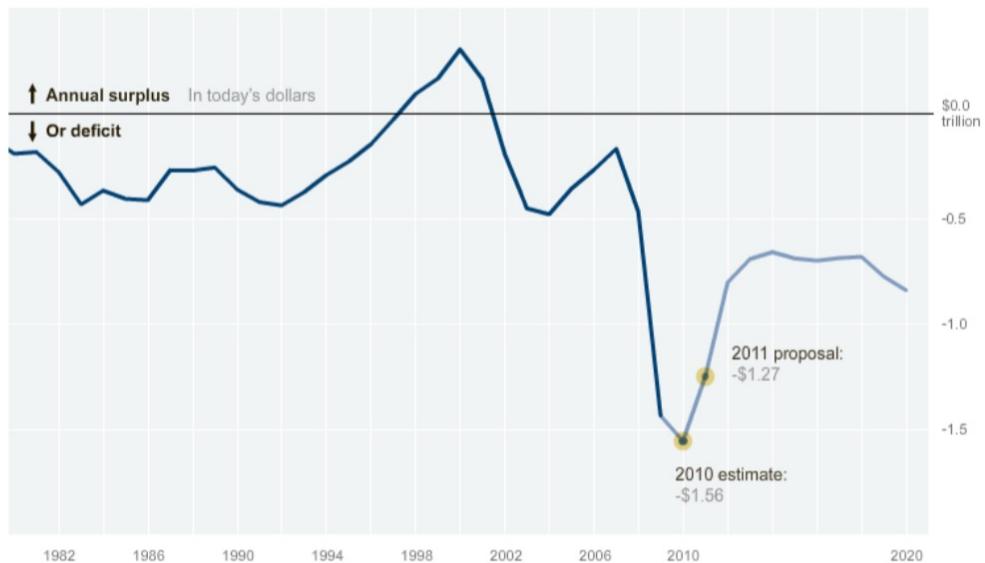
Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT ►

Falling short

President Obama's budget proposal estimates a deficit of \$1.6 trillion for the current fiscal year and \$1.3 trillion in 2011.



By AMANDA COX | Send Feedback

Source: Office of Management and Budget

LINKEDIN SHARE

Figure 6.4 Budget Forecasts, compared with Reality [Cox, 2010].

1. Martini glass: Martini Glass structures begin with an author-driven approach and a default display of the visualization.

- This storytelling structure uses linear author-driven narrative at the beginning, and further extends to reader-driven exploration.
- Consider this example from the New York Times (Refer to Fig. 6.4) [Cox, 2010].

Narrative storytelling

Published: February 2, 2010

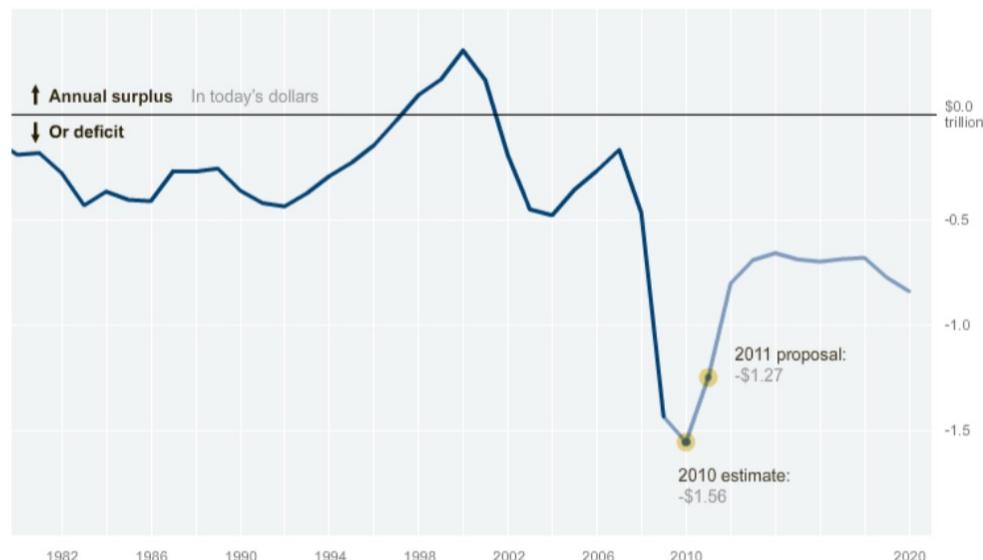
Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT ►

Falling short

President Obama's budget proposal estimates a deficit of \$1.6 trillion for the current fiscal year and \$1.3 trillion in 2011.



By AMANDA COX | Send Feedback

Source: Office of Management and Budget

LINKEDIN SHARE

- Over time, this narrative illustrates the US budget surpluses and deficits from each of the presidential administrations.
- A stepper widget as highlighted in Fig. 6.4 provides linear story navigation to the readers.
- The story through annotations is described in slides one to four.

Figure 6.4 Budget Forecasts, compared with Reality [Cox, 2010].

Narrative storytelling

Published: February 2, 2010

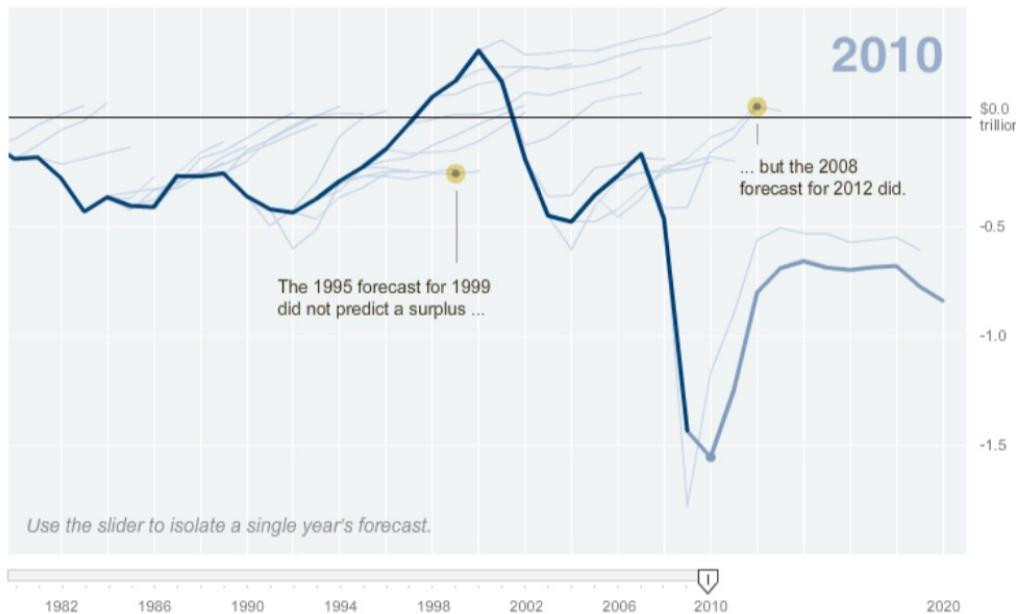
Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT ►

Latest forecast

Today, with a better understanding of the severity of the economic downturn, the deficit situation is much more dire.



By AMANDA COX | [Send Feedback](#)

Source: Office of Management and Budget

[LINKEDIN](#) [SHARE](#)

- The fifth slide (Refer to Fig. 6.5) provides a time slider for the readers to interact with the slider.
- Just like the martini glass approach, the author begins with a linear author-driven narrative, and then opens up the visualization for reader-driven exploration.

Figure 6.5 Budget Forecasts, compared with Reality [Cox, 2010].

Narrative storytelling

2. Interactive slideshow:

- The interactive slideshow consists of a conventional slideshow which allows interactivity with various parts of the presentation before allowing a story to reach its maturity state.
- For example, you can create a dashboard with interactive data visualization using Tableau and embed these in your PowerPoint presentation.

Narrative storytelling

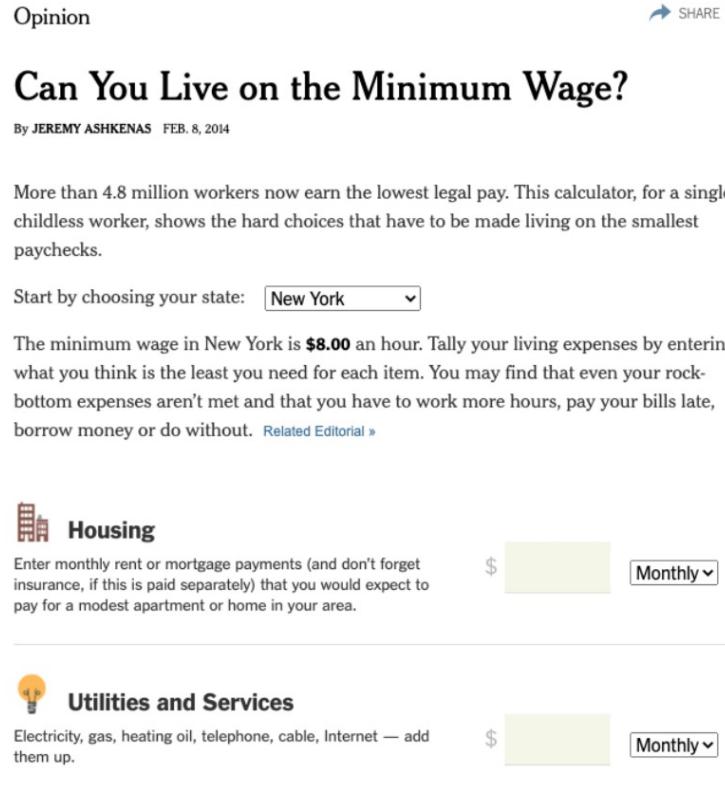


Figure 6.6 Minimum wage to live [Jeremy, 2014].

Source: <https://www.nytimes.com/interactive/2014/02/09/opinion/minimum-wage.html>

3. Drill down story:

- This is a reader-driven approach in which we present a general theme and allow user interactivity with certain parts of the visualization.
- A New York Times example is shown in Fig. 6.6 [Ashkenas, 2014].
- In this case, the reader can explore the story through his own preferences by selecting appropriate dropdown values.

Dimensions of Narrative Storytelling

- Segal and Heer (2010) researched and presented dimensions for visual storytelling as follows
 1. Genres,
 2. Visual Narrative tactics, and
 3. Narrative framework.

Genres

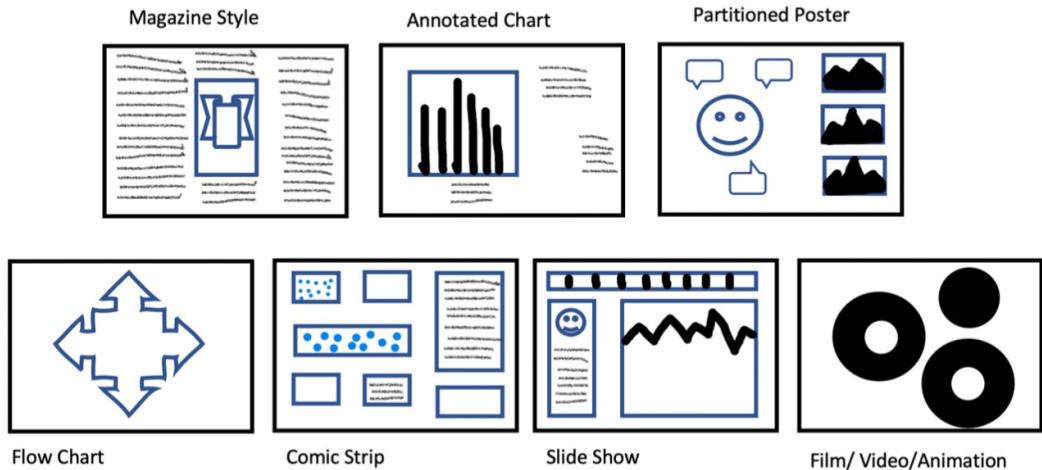


Figure 6.7 Genres of narrative visualization by Segel and Heer.

Source: https://www.researchgate.net/figure/The-figure-shows-the-seven-genres-of-narrative-visualization-presented-by-Segal-and-Heer_fig19_323777213

- The purpose of the genres is to tell a story.
- There are seven genres of narrative visualization, such as Magazine style, Annotated chart, Partitioned poster, Flow chart, Comic strip, Slide show, and Video
 1. Distinct visual scenes
 2. Ordering of the visual elements

Visual Narrative Tactics

- Techniques for organizing visual media to produce an informative narrative experience are called visual narrative tactics.
- This can be classified as follows:
 1. **Visual structuring:** This helps orient the reader to the narrative.
 2. **Highlighting:** It draws the reader's attention to important aspects of visualization.
 3. **Transition guidance:** Techniques of transitioning within or between visual scenes without confusing the readers.

Narrative Framework

- Narrative frameworks reduce uncertainty, connect data with context, and describe specific interpretations. This involves three components:

1. Ordering:

- How we can organize the story.
- It helps figure out ways to arrange the path readers take through the visualization.
 1. **Linear:** It can be linear and predetermined by the author.
 2. **Random access:** There will be no suggested path, so readers will have random access.
 3. **User directed:** User selects a path from multiple alternatives.

Narrative Framework

2. Interactivity:

- Engage the user in the story by providing different ways to manipulate visualization.
- For example, filtering, searching, and navigating.
- By allowing readers to interact with the data, they can quickly change the view, making analysis quicker and more effective.

3. Messaging:

- Use techniques such as labelling, captions, headlines, and annotations to provide commentary for readers.
- Context to the narrative is added by using captions, labels, or other annotations.
- There are two types of annotations: Observational and Additive.

Narrative Framework

a) Observational:

- These annotations are chart creator's view-point that give context to a data value that is depicted in the visualization.
- One example would be, highlighting an outlier or providing more information as we hover over visual elements.

a) Additive:

- Additive annotations add external context to a visual representation that cannot be clearly depicted through the data.
- This is key information to the topic or critical to the understanding of the data, such as providing background on the events or actions.

Narrative Framework

- Given below are links to a few interesting narrative visualizations which we can explore and learn from.
1. [Interesting US elections analysis at https://projects.fivethirty-eight.com/2018-midterm-election-forecast/house/](https://projects.fivethirty-eight.com/2018-midterm-election-forecast/house/)
 2. [How birth year influences the political view, An interesting data story – https://www.nytimes.com/interactive/2014/07/08/up-shot/how-the-year-you-were-born-influences-your-politics.html](https://www.nytimes.com/interactive/2014/07/08/up-shot/how-the-year-you-were-born-influences-your-politics.html)
 3. [A narrative with visuals on machine learning built using R3D3 package – http://www.r2d3.us/visual-intro-to-machine-learning-part-1/](http://www.r2d3.us/visual-intro-to-machine-learning-part-1/)

Data Story Types

- The narrative data story helps provide context and enables us with options to start a discussion about actions or drive home key points to the audience.
- Ben Jones from Tableau put together the following types of data story as a starter for us to build our data story [Jones, 2014].
- There are seven basic types of data stories which we can use to tease stories from the data.
- These data stories are quantitative visual narratives, or a series of facts shown in a linear way through data visualization.

Data Story Types

- We will explore different data story types using the story points feature in Tableau.
- We will use an interesting dataset about Crimes in the Open Govt Data Platform India portal.¹
- This dataset consists of several crimes that happened in India from 2001 to 2014.
- It covers major IPC (Indian Penal Code) sections such as Murder, Rape, Auto theft, Dacoity etc.
- Assume that we are building a presentation for our stakeholders interested in learning more about the problem statement and looking through opportunities to partner with our new fighting against crime initiative.

¹Source:<https://data.gov.in/>

Data Description

Table 6.1 Crimes in India (2001–2014)

Variable	Variable Type	Description
State	Categorical	Indian States
District	Categorical	Indian Districts
Year	Date Time (Year)	Year of Crime
NumberOfCrimes	Numerical	Total number of crimes
CrimeType	Categorical	Type of Crime – 20 major categories
CrimeGroup	Categorical	Major crimes have been grouped under the following categories: 1. Homicide 2. Crime against women 3. Kidnapping and abduction 4. Property stolen 5. Serious fraud 6. Others
Region	Categorical	Indian zones

- Seven Data Story types are:
 1. Change Over Time
 2. Drilling Down
 3. Zooming Out
 4. Contrast
 5. Intersection
 6. Factors
 7. Outliers

Change Over Time – Story Point 1

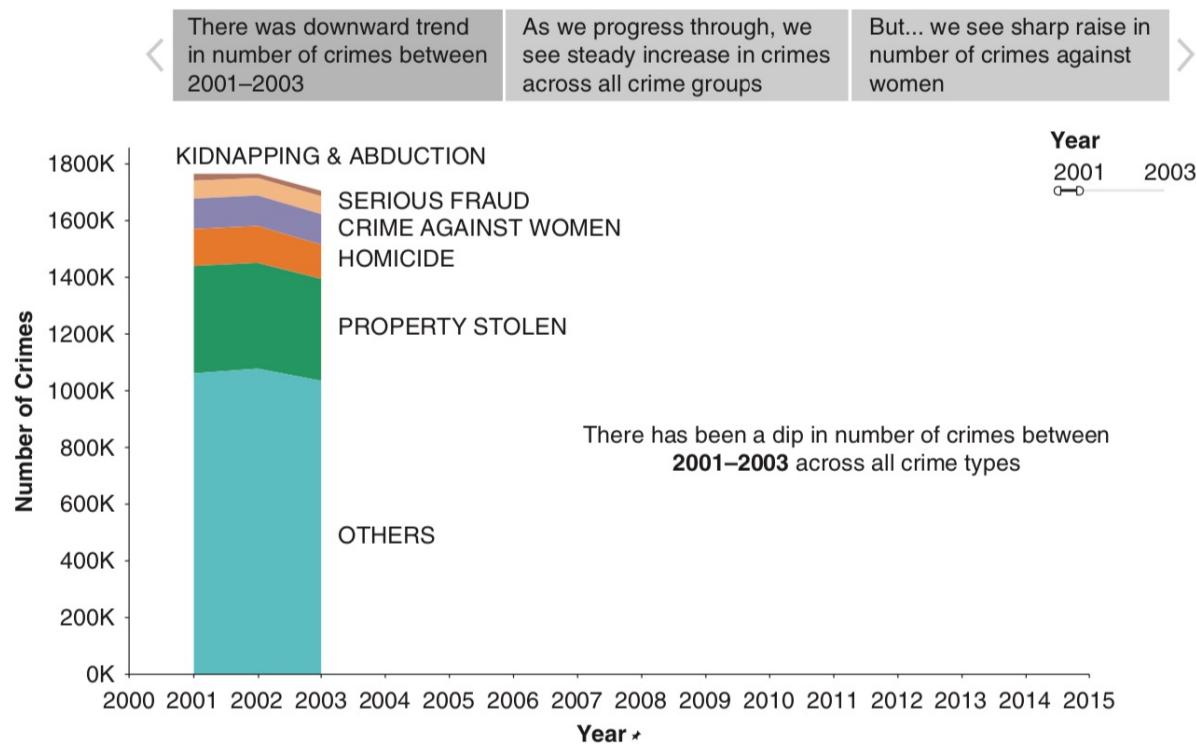
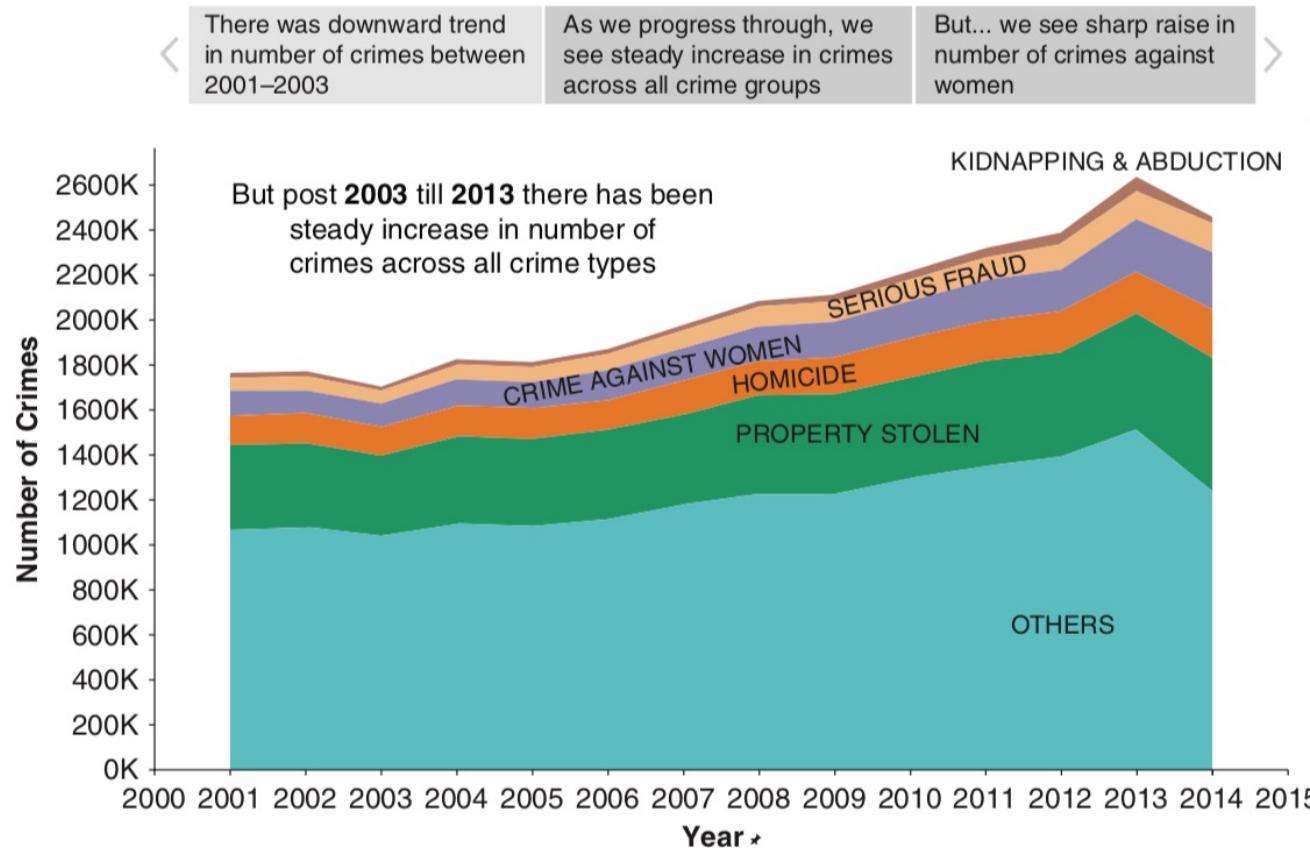


Figure 6.8 Change over time story – story point 1.

- In this case we are looking at how a particular factor has performed.
- We can see how the number of crimes changed over time between 2001 and 2014, and what story they tell.
- A central theme in this story will be “Crimes against Women”.
- Between 2001 and 2003, the number of crimes in India decreased.

Change Over Time – Story Point 2



- As India progressed through 2004 to 2014 in terms of economic growth, gaining diplomatic clout, digitization, the arts, and sports.
- We also saw a steady increase in crimes across all crime groups.

Figure 6.9 Change over time story – story point 2.

Change Over Time – Story Point 3

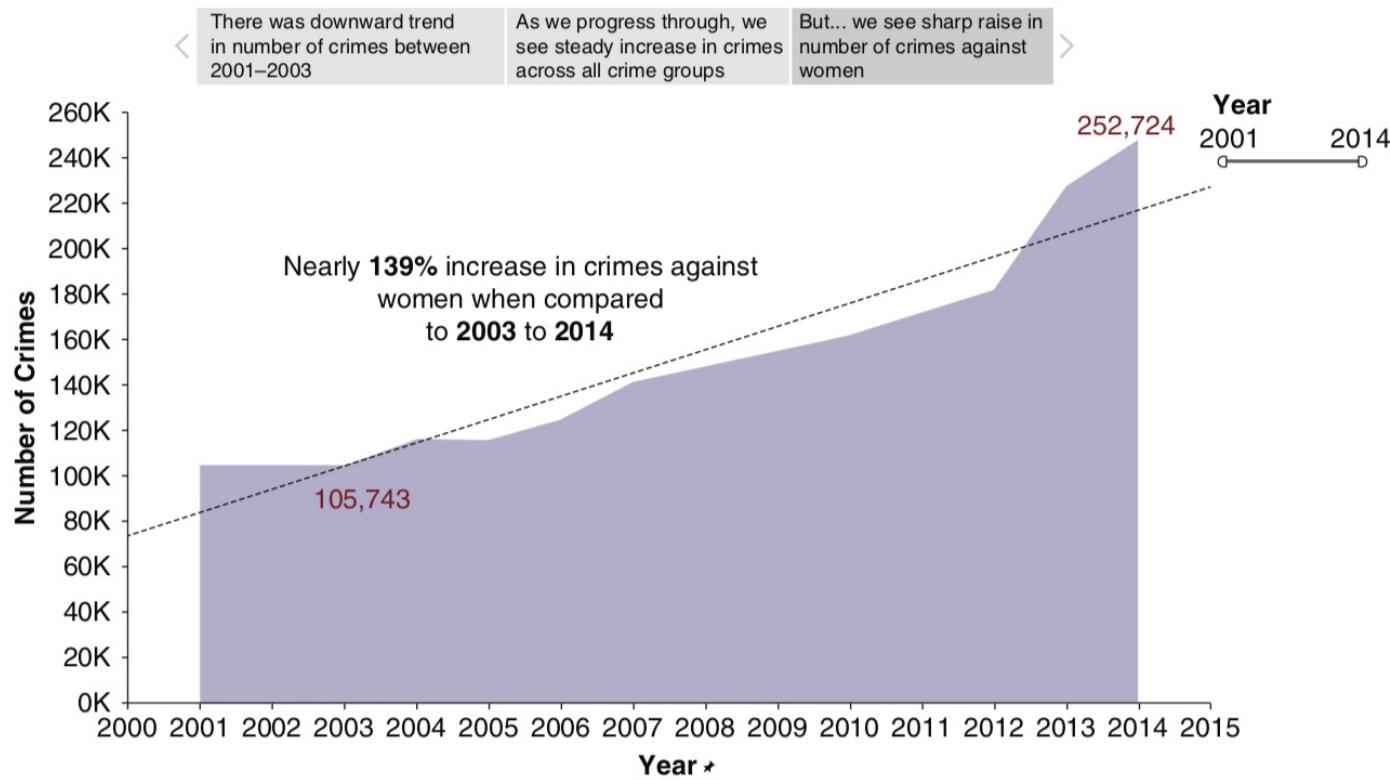


Figure 6.10 Change over time story – story point 3.

- When we examine the “Crimes against Women”, we see a sharp rise in the numbers.
- From 2003 to 2014, crimes against women more than doubled.

Drilling Down – Story Point 1

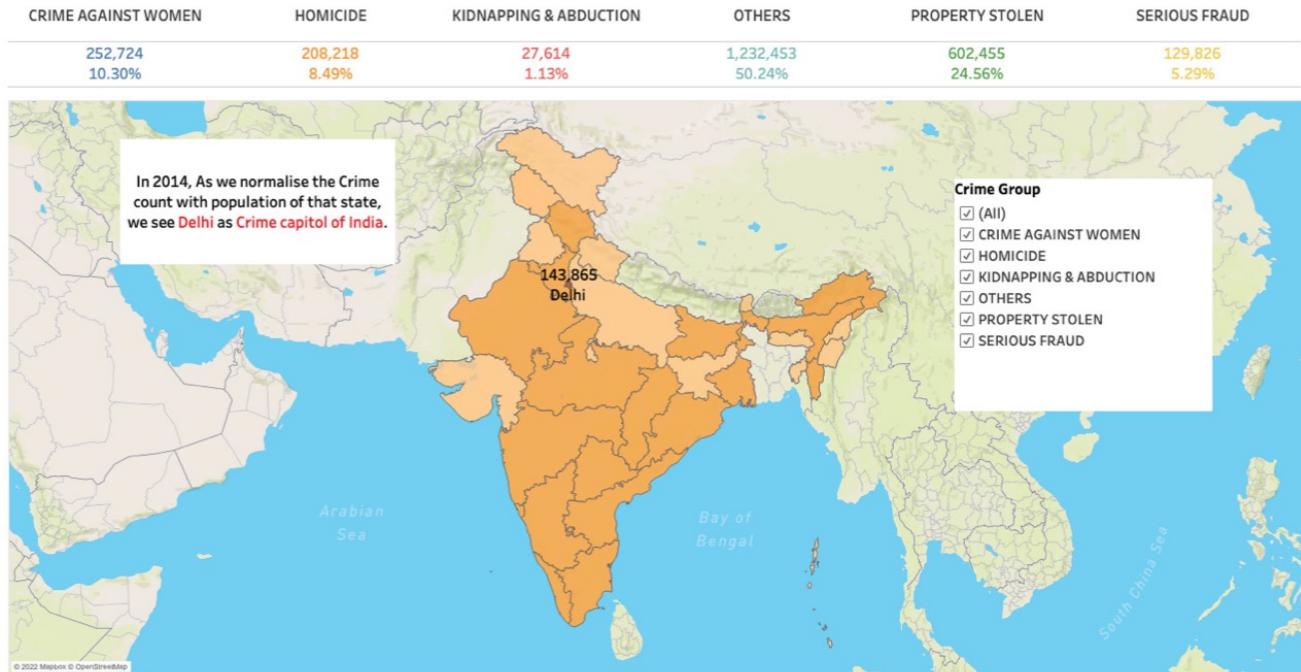


Figure 6.11 Drill down story – story point 1.

- If we have **hierarchical data**, we can see how a particular dimension affects the system.
- Figure 6.11 shows that even though we have crime numbers high in Madhya Pradesh, Maharashtra, and Utter Pradesh.
- When we normalise these crime numbers with the state's population, we get Delhi with highest crime rates.

Drilling Down - Story Point 2

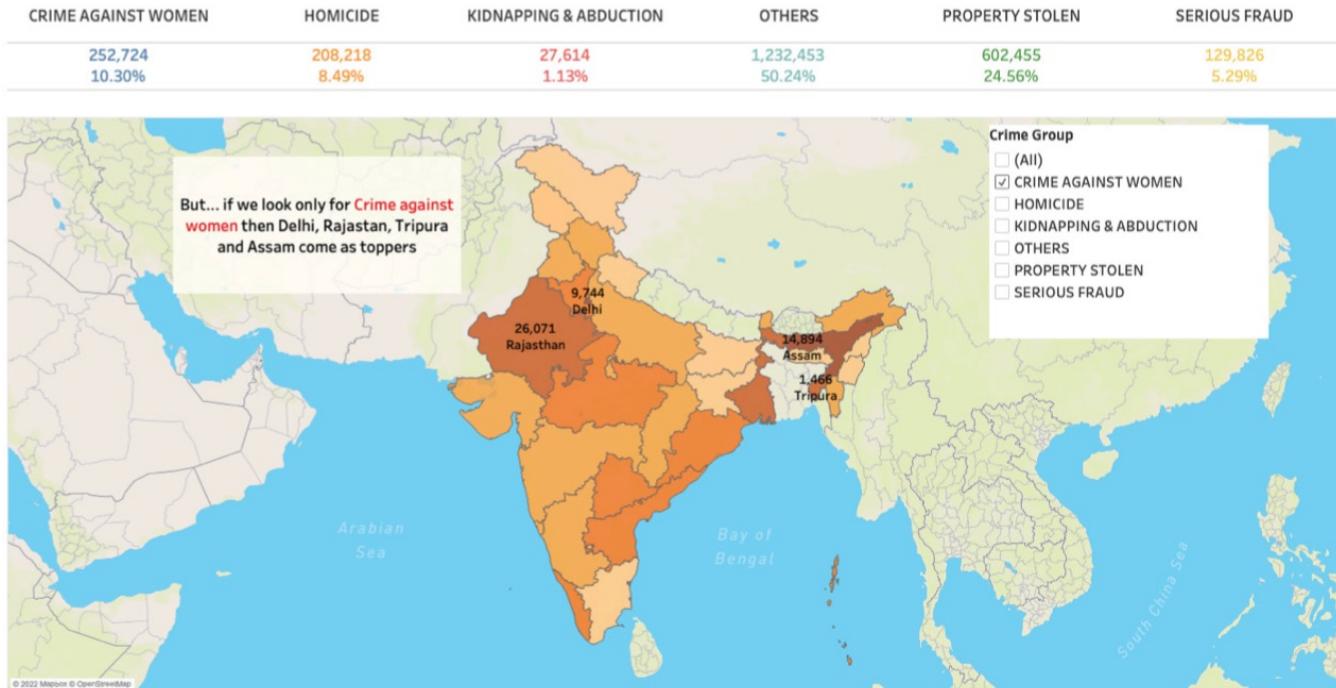
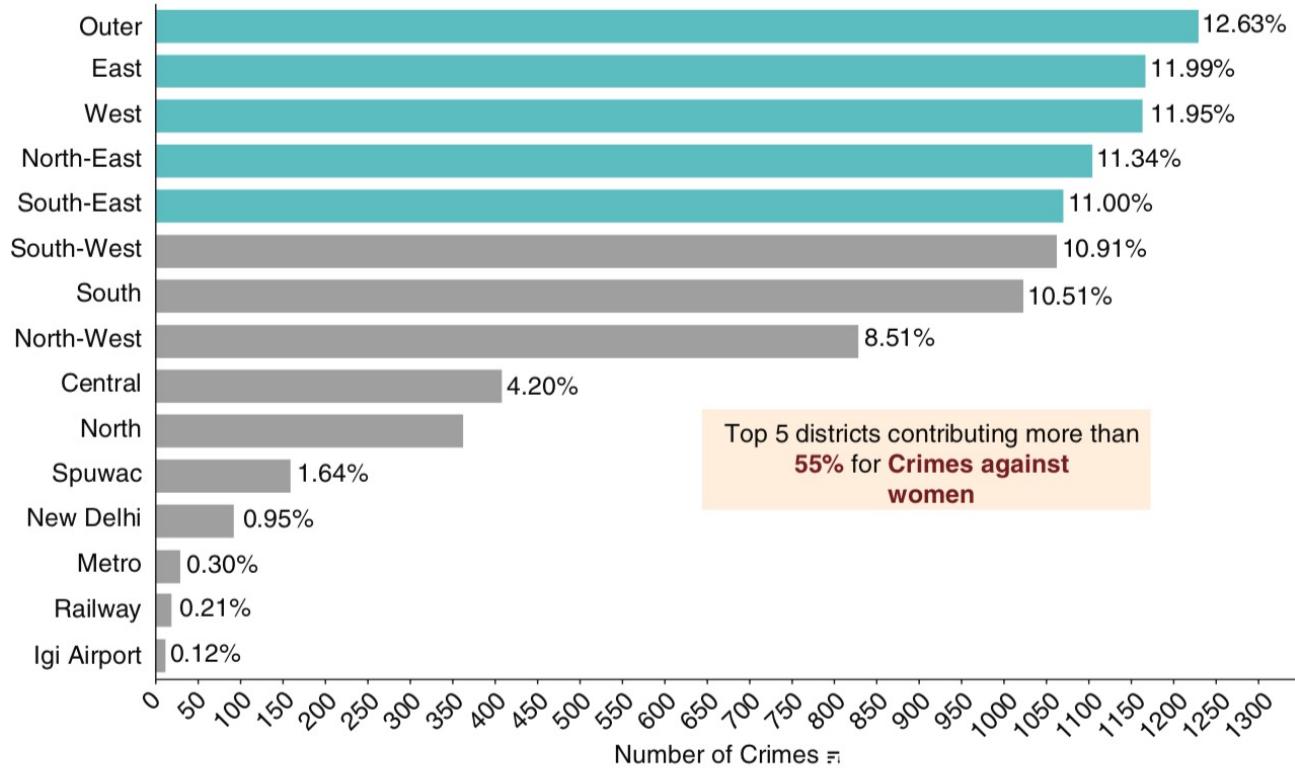


Figure 6.12 Drill down story – story point 2.

- When we drill down to see where the crimes against women are most prevalent, we see even here Delhi is in the top quadrant along with Rajasthan, Tripura, and Assam

Drilling Down- Story Point 3



- Outer, East, West, Northeast, and Southeast are the top five districts for crimes and contribute to more than 55% of crimes against women in Delhi.

Figure 6.13 Drill down story – story point 3.

Drilling Down - Story Point 4

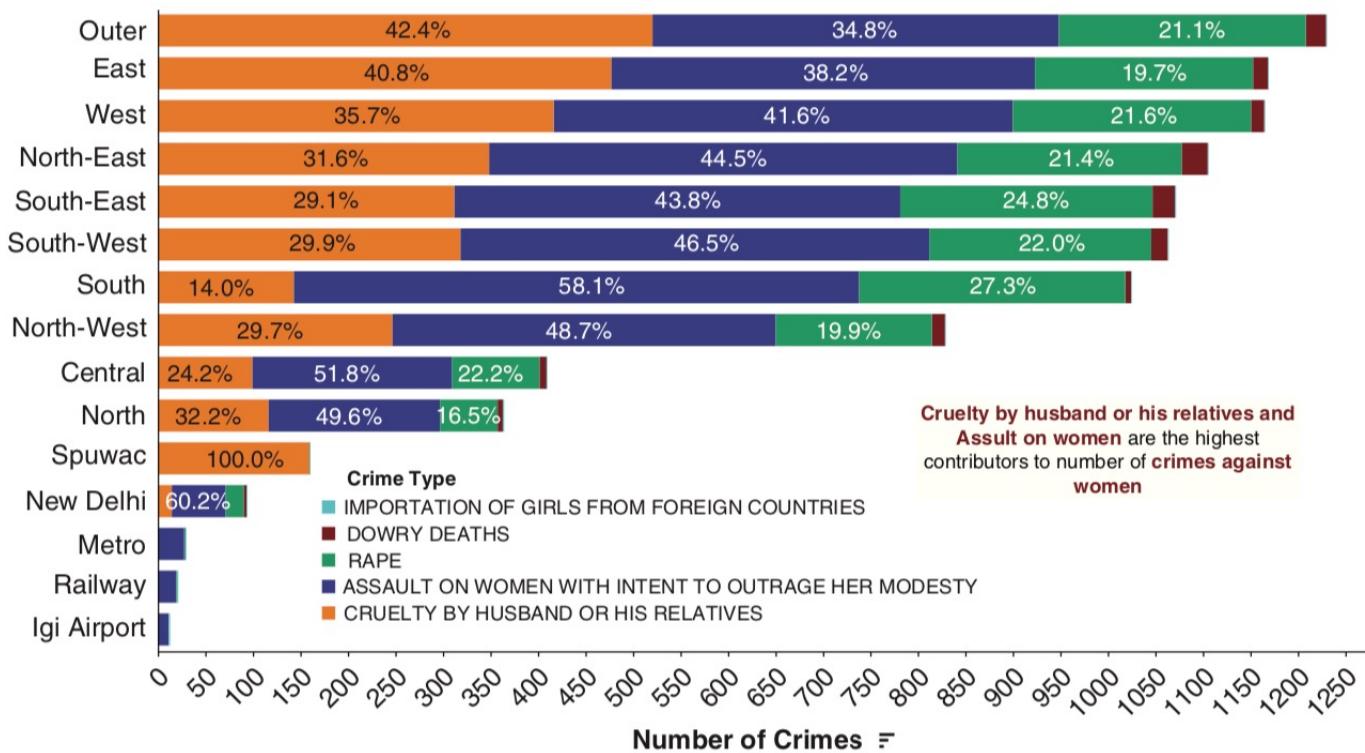


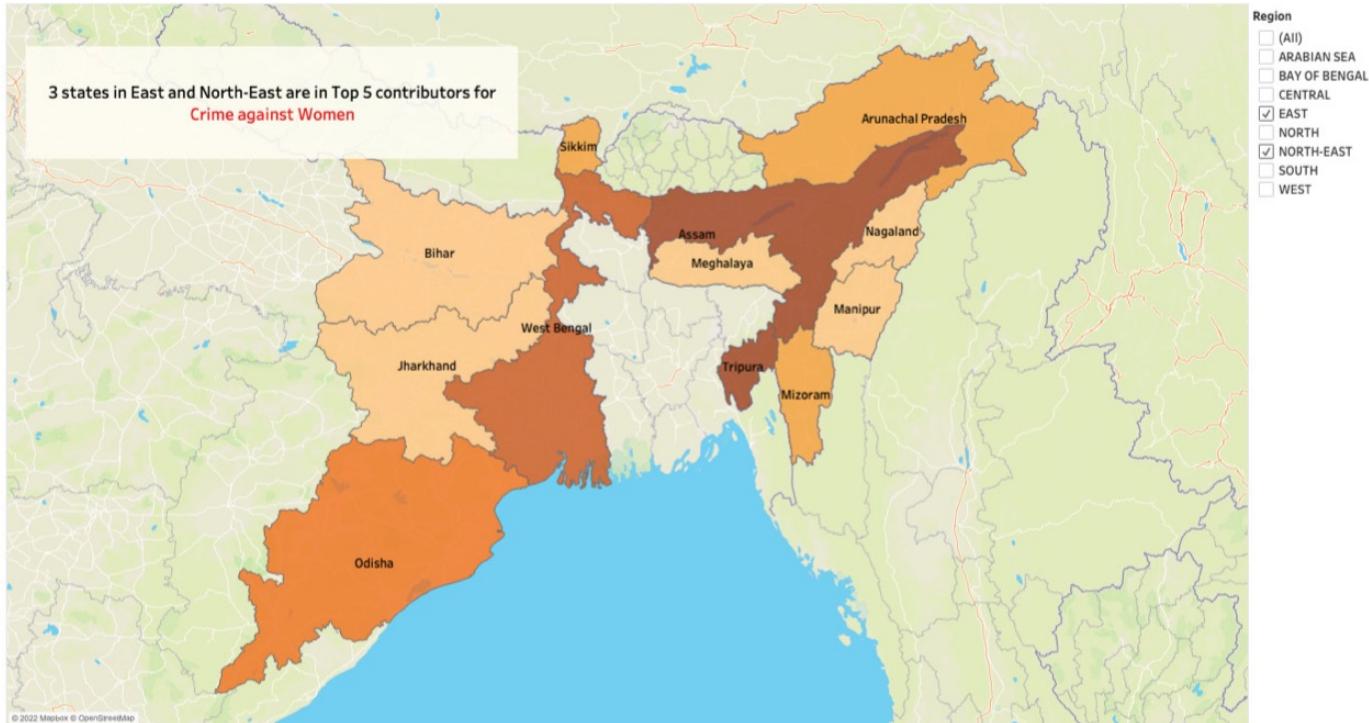
Figure 6.14 Drill down story – story point 4.

- We can see (Refer to Fig. 6.14) that
 1. “Cruelty by Husband or his relatives” – IPC Section 498-A3 and
 2. “Assault on women with intent to outrage her modesty” – IPC section 354are the major contributors to crimes against women.

Zooming Out

- Think small and zoom out is a story type which helps us tell stories about how small details affect the whole.
- Zoom out is a kind of opposite of drilling down but there is a slight difference.
- In the zoom out data story type, we focus our attention on one particular area of data and comparing it to the rest of the data or we can start broadly and focus on a particular area of interest.
- Using this story type helps us shift the perspective of the audience.

Zooming Out- Story Point 1



- Refer to Fig. 6.15.
- In the Top-5 states with the highest crimes against women, 3 states are from East and North-East India.

Figure 6.15 Zoom out – story point 1.

Zooming Out- Story Point 2

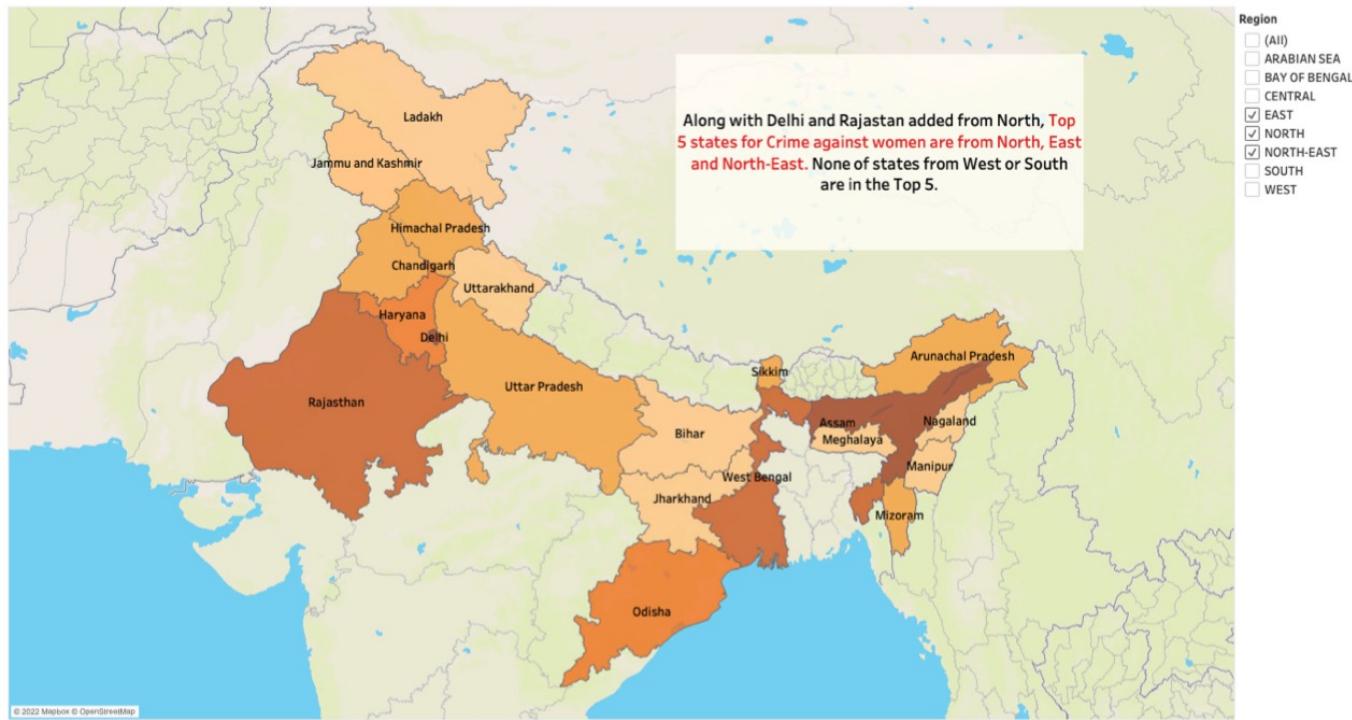


Figure 6.16 Zoom out – story point 2.

- Fig. 6.16, we see that none of the western or southern states features at the top when it comes to contributing towards crimes against women.
- It would be interesting to find out why crime rate for crimes against women is low in these regions of India.
- We can check how these numbers fare by considering the socio-economic parameters from these regions.

Zooming Out- Story Point 3



Figure 6.17 Zoom out – story point 3.

- Refer to Fig. 6.17.
- If we keep our focus on regions of India and look at all the crime groups put together, we see that southern state – Kerala also featuring in top states contributing towards crimes in India.

Contrast - Story Point 1

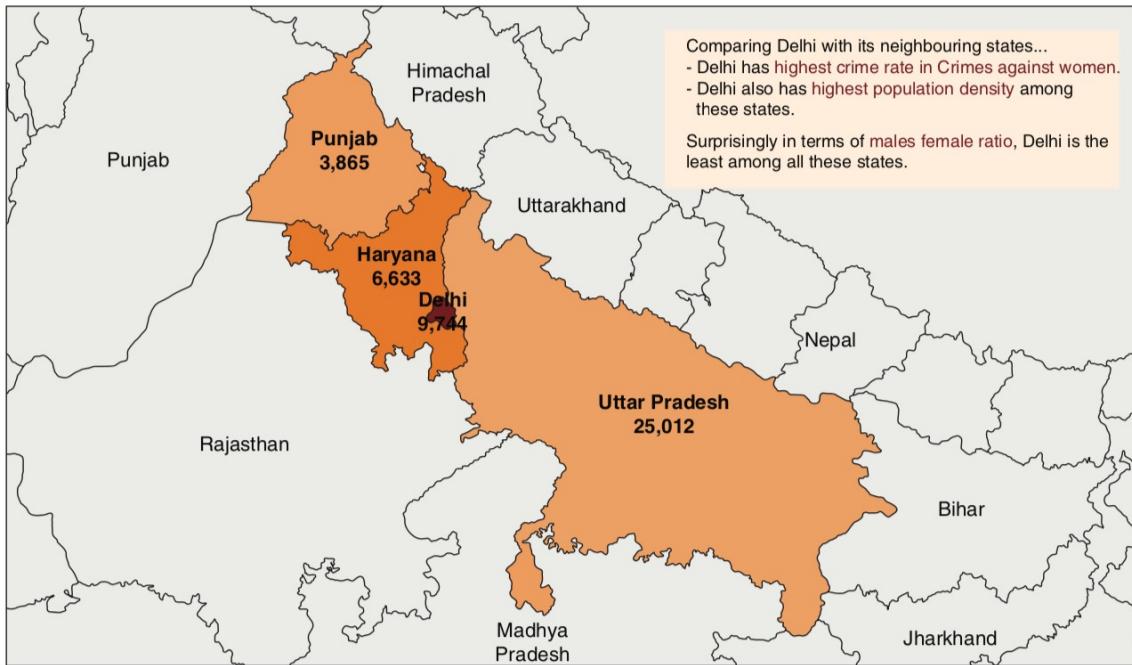


Figure 6.18 Contrast – story point 1.

- This storytelling type will be useful when we want to compare the difference between two categories of data or compare the progress of one group with another.
- we compare Delhi with its neighbouring states – Haryana, Punjab, and Utter Pradesh.
- Delhi has highest crime rate in Crimes against women.
- Delhi also has highest population density as compared with its neighbouring states.
- Contrastingly if we check for sex ratio⁷ for these states, Delhi has the least female to male ratio at 868 females for every 1000 males.

Contrast - Story Point 2

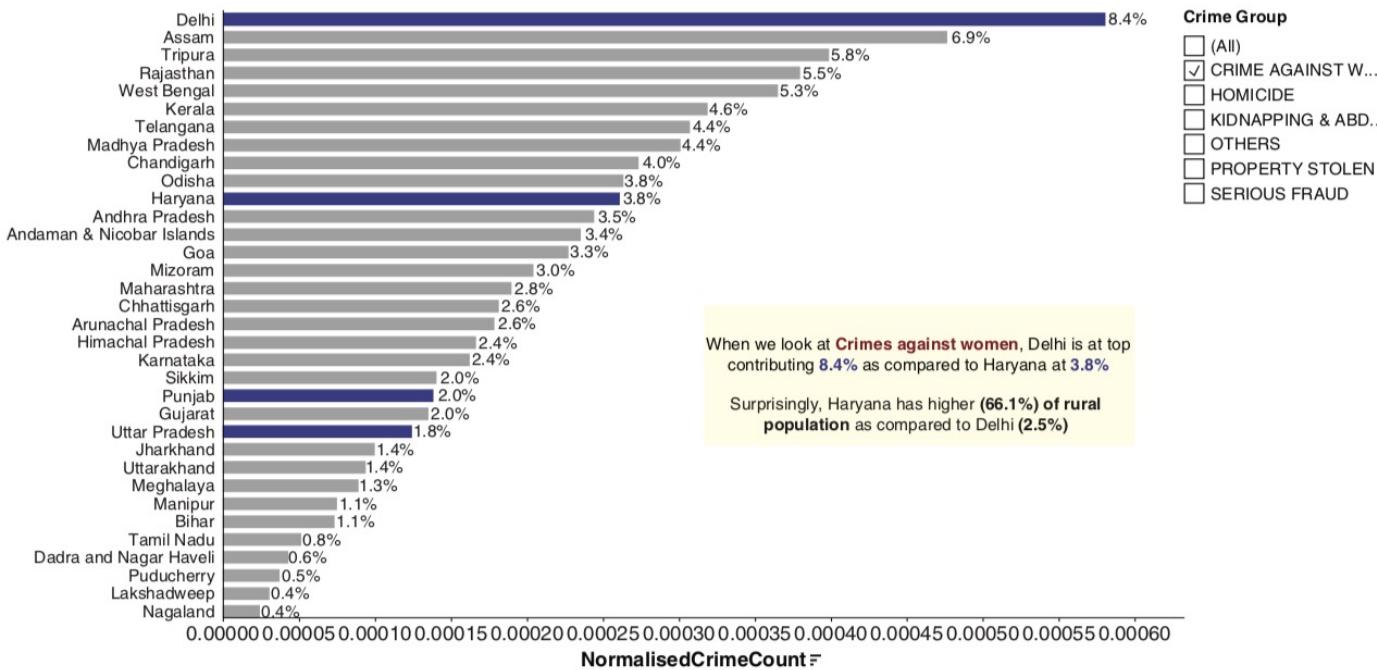


Figure 6.19 Contrast – story point 2.

- Overall Delhi is contributing 8.4% of crimes against women.
- Surprisingly even with higher rural population as compared to Delhi, Haryana is having lesser crime rate (3.8%) for crimes against women.

Contrast- Story Point 3

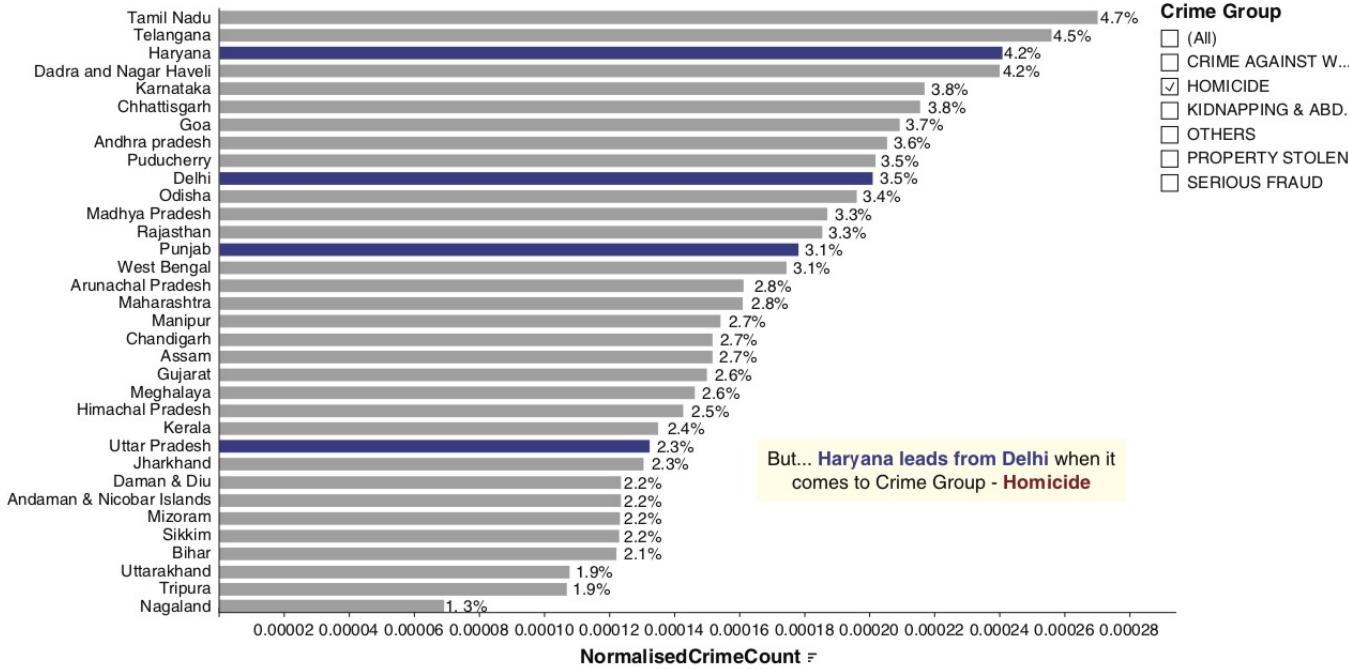


Figure 6.20 Contrast – story point 3.

- If we look at other crime group like Homicide, we can see Haryana leads from Delhi in this crime group category.
- When we present our viewpoint using contrast story type, we will be able to discuss:
 1. Reasons for these differences.
 2. What are the things done right by one group which we can apply to another?

Intersection - Story Point 1

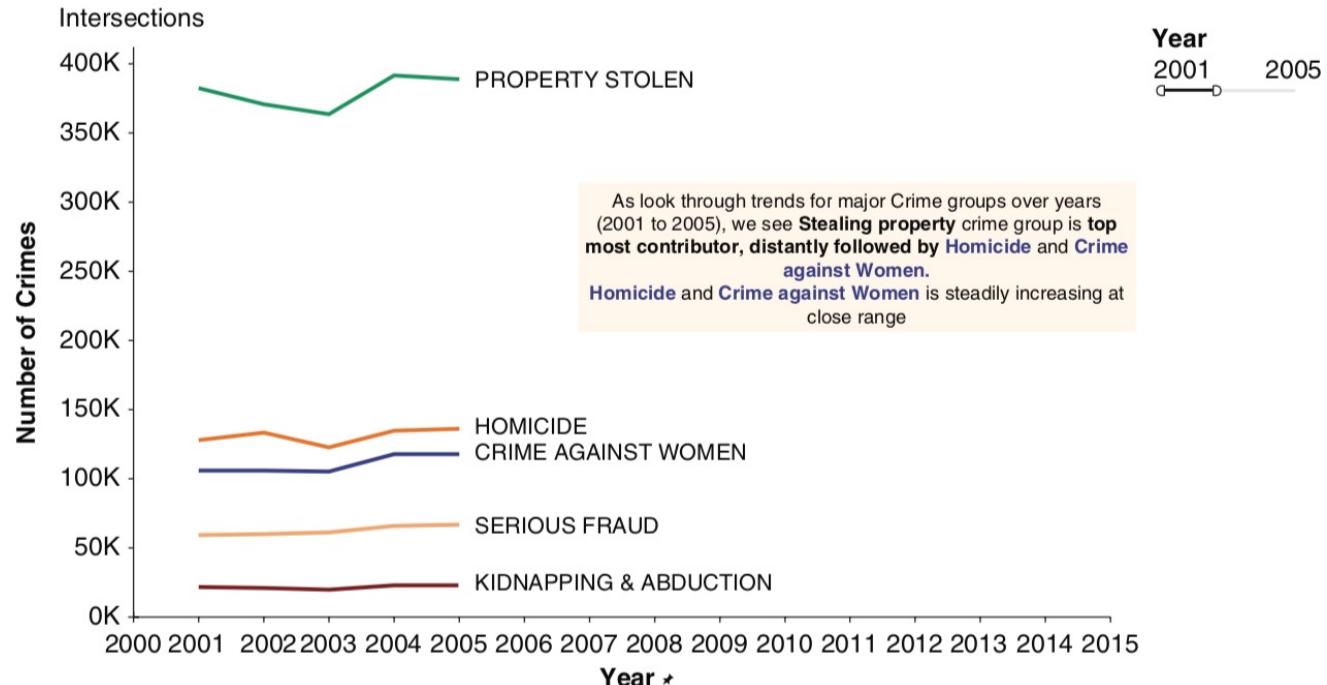


Figure 6.21 Intersections – story point 1.

- Intersections story type helps us explore reasons why one category overtakes another.
- Refer to Fig. 6.21, when we look at the trend between 2001 and 2005 across major crime groups.
- We can see that “Stealing Property” is one of the major contributors to overall crime numbers.
- Homicide and crime against women take distant second and third positions, respectively.

Intersection - Story Point 2

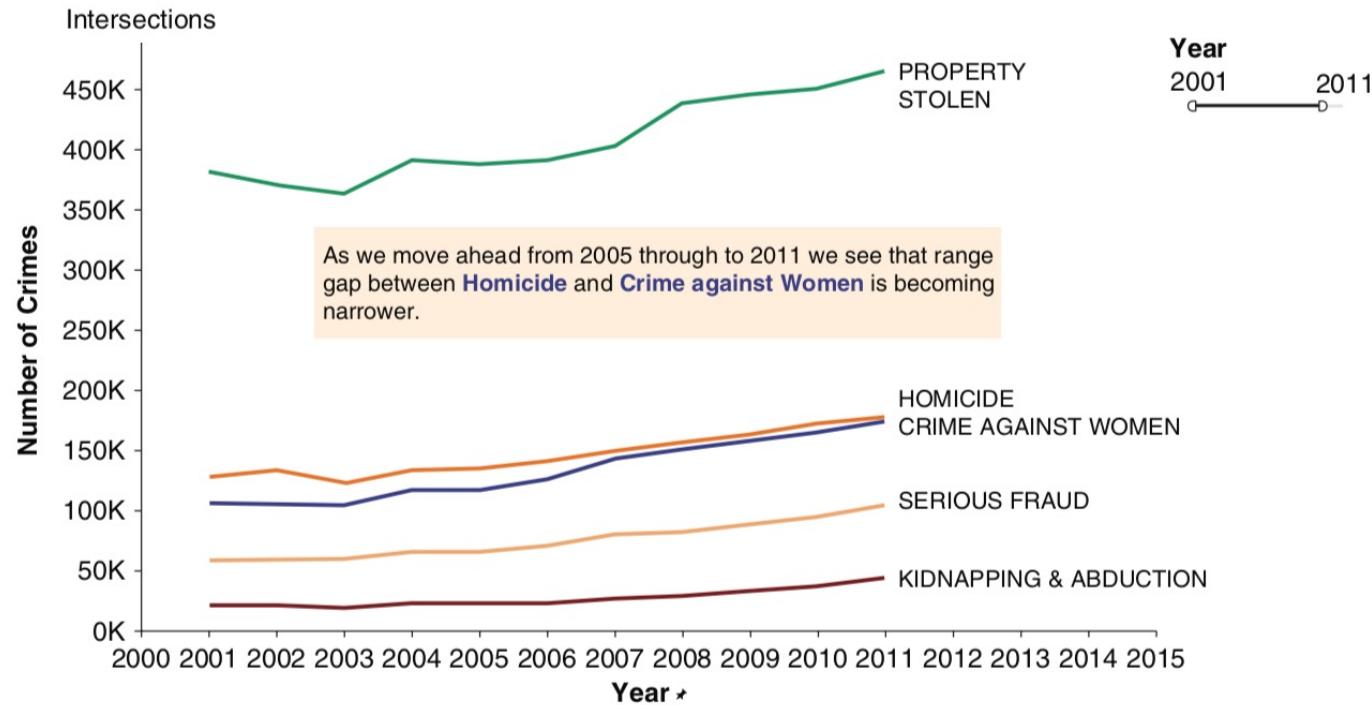
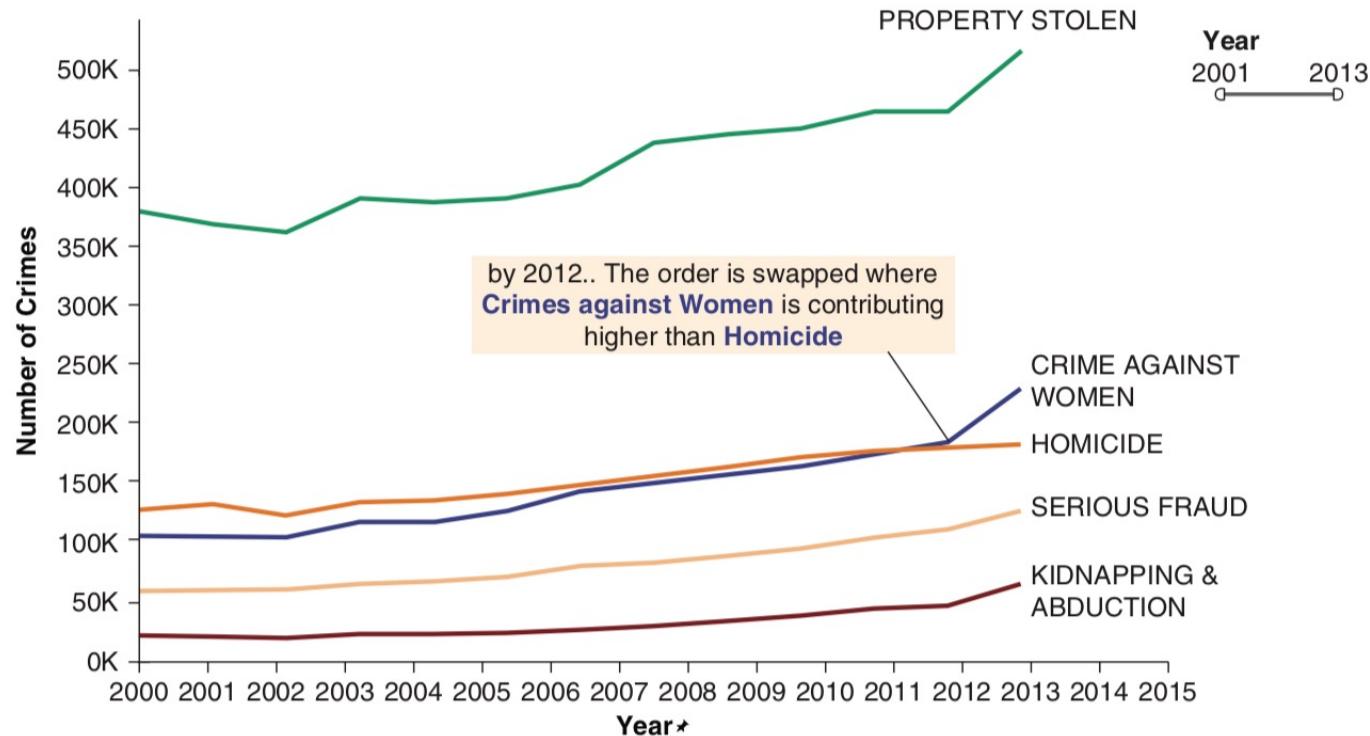


Figure 6.22 Intersections – story point 2.

- Refer to Fig. 6.22
- As we move through 2005 to 2011, we see that the gap between homicide and crime against women crime groups is reducing.

Intersection - Story Point 3



- Refer to Fig. 6.23
- By 2012, the order is swapped where crimes against women takes second place and homicide steadily moves to third place.

Figure 6.23 Intersections – story point 3.

Intersection- Story Point 4

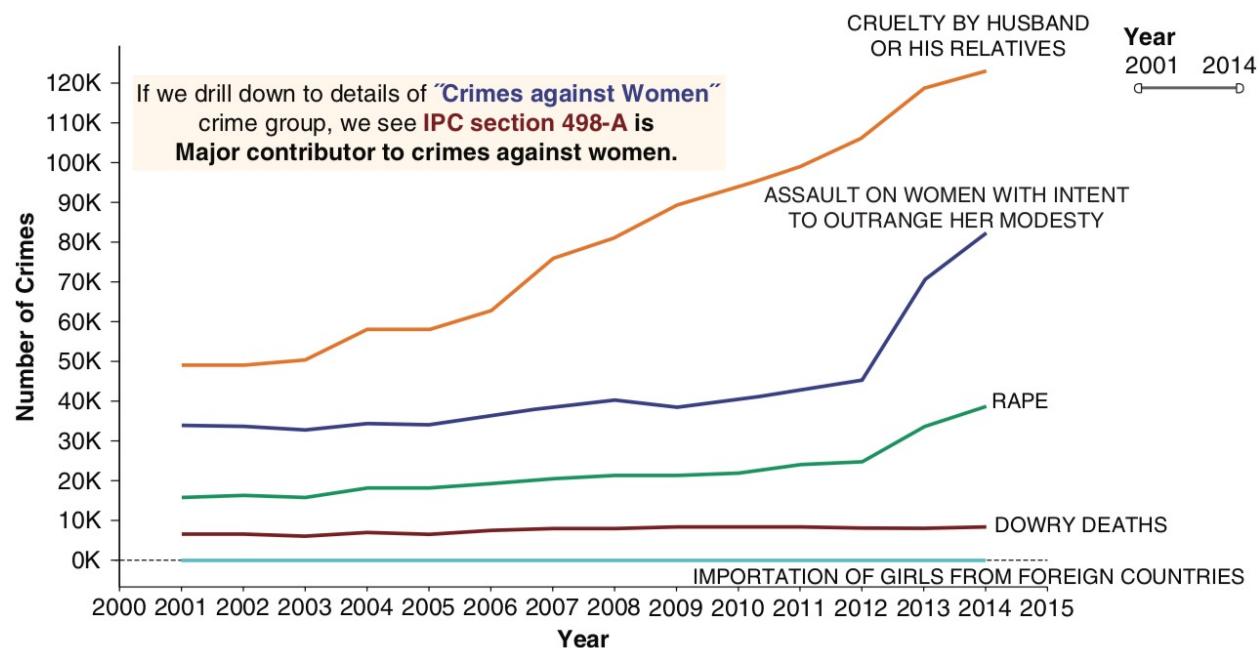


Figure 6.24 Intersections – story point 4.

- In Fig. 6.17 we explore detailed trend across what crime types contribute significantly towards “Crimes against women”.
- We see IPC section 498-A, Cruelty by Husband, or his relatives, is the highest contributor.
- It also tells us the disturbing fact that crimes against women, especially rape and molestation, are on the rise in the country.

Factors - Story Point 1

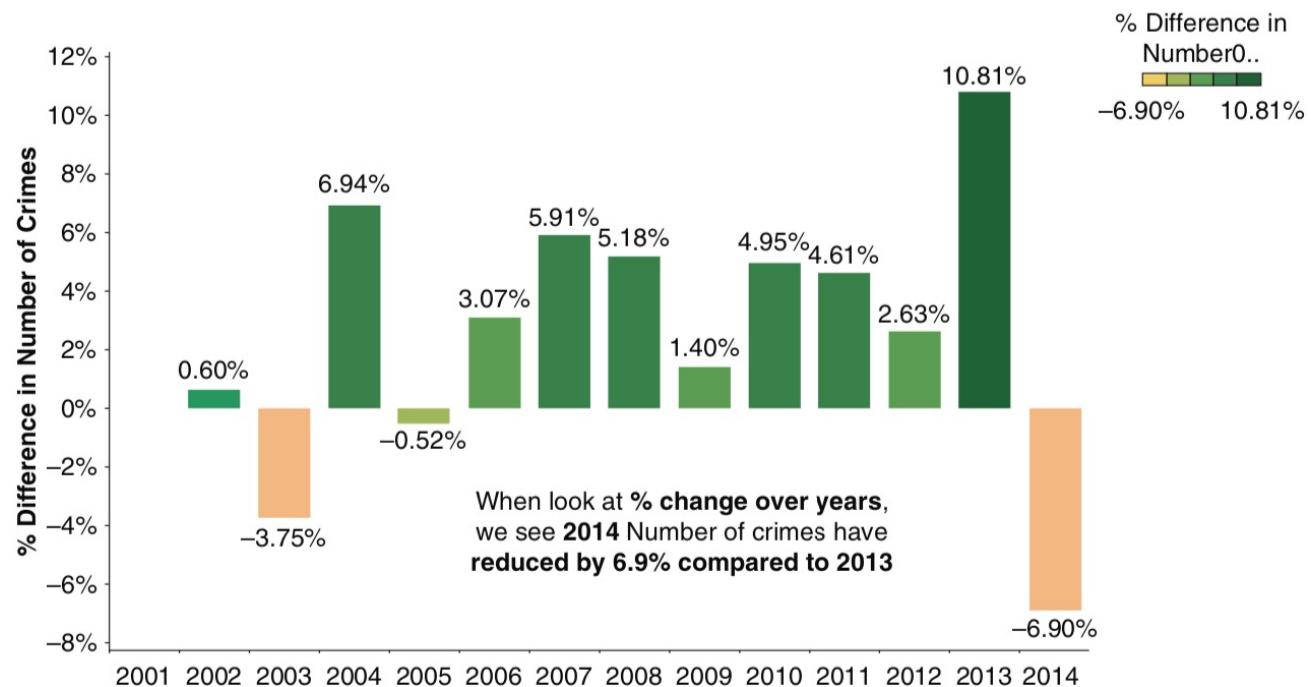


Figure 6.25 Factors - story point 1.

- Factor data stories can help us while checking for correlation or causation between metrics.
- We consider one main metric and look for the factors that have the greatest influence.
- In Fig. 6.25, we look at the percentage change in the overall crime numbers trend between 2001 and 2014.
- Crimes have been down by 6.9% in 2014 compared to 2013.

Factors - Story Point 2

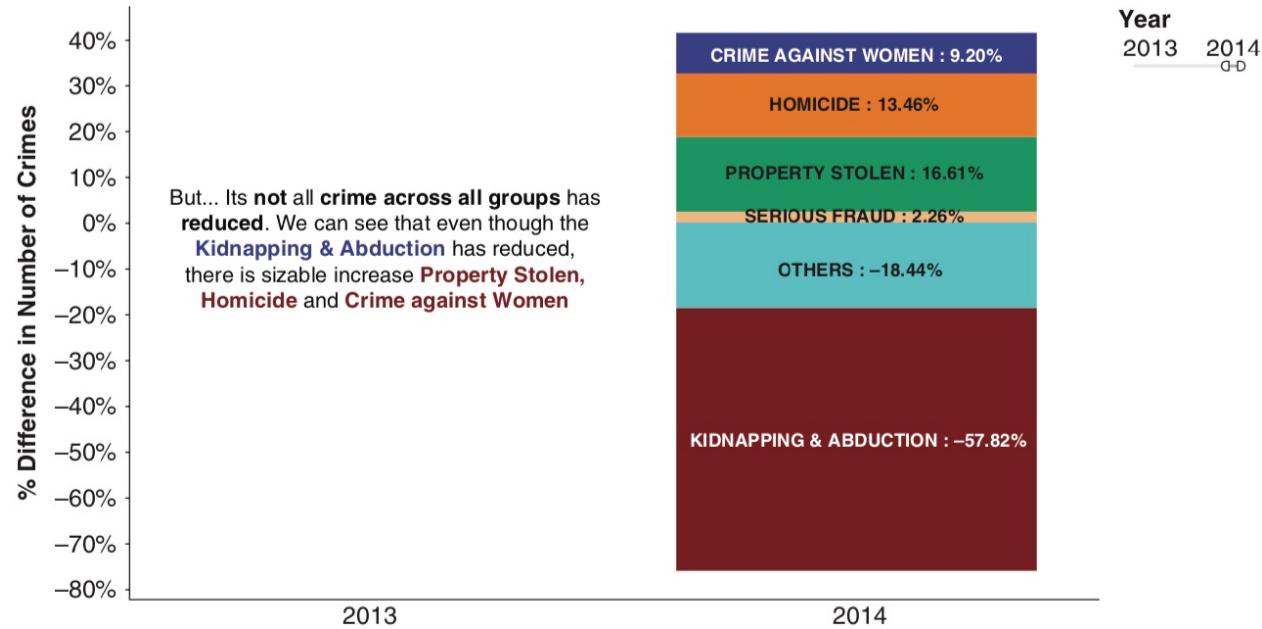


Figure 6.26 Factors – story point 2.

- Refer to Fig. 6.17.
- But if we check on factors like detailed break-up of this trend for each crime group in 2014, we can see that not all crime groups have a negative trend.
- Even though kidnapping and abduction has reduced, there is a sizable increase in property stolen, homicide and crime against women.

Outliers - Story Point 1

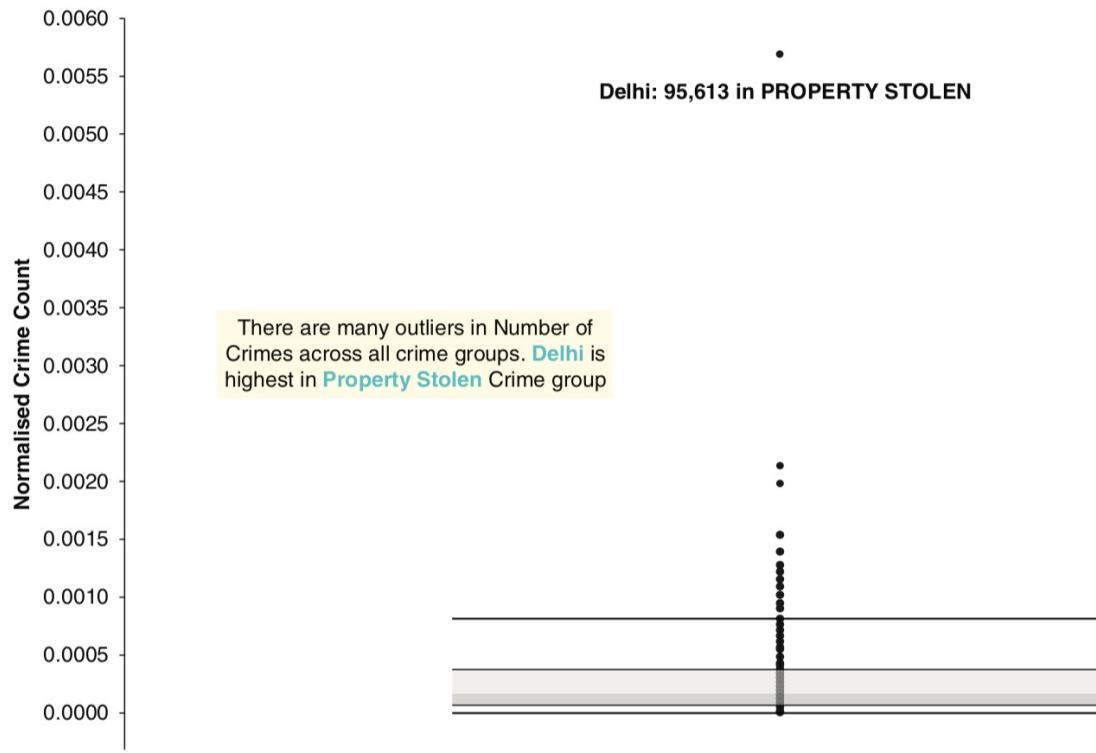


Figure 6.27 Outliers – story point 1.

- This story type is all about pointing out specific areas where things are substantially different.
- When we check the outliers in the number of crimes data for 2014, Delhi with property stolen crime group has the highest value.
- Jammu & Kashmir has the least value for crime against women.
- This indicates probable missing data from the dataset as Jammu & Kashmir has zero crime against women in 2014, which cannot be true.

Outliers- Story Point 2

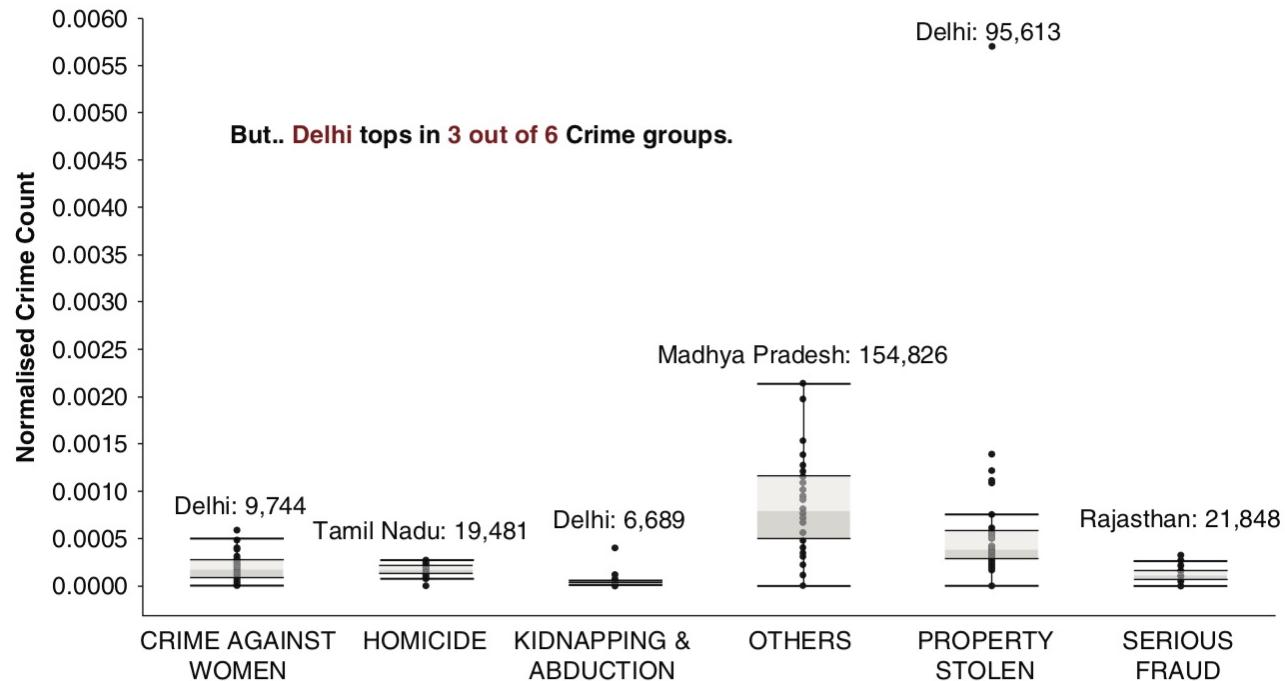


Figure 6.28 Outliers – story point 2.

- If we look for outliers in crime numbers across all crime groups, we get different states contributing the highest outlier value.
- Out of six crime groups, Delhi has highest crime rate in three crime groups out of six.
- Interestingly, Tamil Nadu has the highest number of “Homicide” cases.
- We can explore this further to check if these are due to the caste-based murders which peaked in southern districts of Tamil Nadu in 2014 [Janardhan, 2015].

Data Story Type

- These seven different story types can be used as thought starters.
- We can use one or multiple of these to build our narrative.
- When we have huge dataset and we are overwhelmed by where to start, these story types can help us explore, uncover, and build interesting stories.
- In any kind of dataset, there is always a story; it is our job as data analysts to find and communicate those stories.

Analytics Dashboard

- The different story types help us explore the data and present our results to our audience.
- To provide an interface to a business or operational audience who would need to interact with the data and understand the analytics process, we need an analytics dashboard.
- Analytics dashboard is built to present the inner workings of our overall analytical solution, which can even be used by a non-technical user.
- It provides a holistic view of all business Key Performance Indicators (KPI).
- Using interactive analytical dashboards, business and operations users can track and monitor their KPIs.

Dataset Description

- Public domain dataset [heartdiseases.csv] about heart diseases from Kaggle is taken to build the analytics dashboard.
- We will use Plotly dash and python SHAP.
- This curated dataset combines different datasets with over 11 common features.
- It is one of the largest heart disease datasets available for research purposes.
- Cardiovascular diseases (CVDs) are the topmost disease-causing death all over the world.
- About 17.9 million people die from it every year, or about 32% of all deaths worldwide.
- CVDs include conditions of the heart and blood vessels such as coronary heart disease, cerebrovascular disease, rheumatic heart disease, and other conditions.

Dataset Description

- We built an ML model (Random Forest) [Kumar, 2021] to predict the risk of heart disease in a patient.
- Based on unseen data (validation data), our model predicts the risk of heart disease with 95% accuracy.
- Since this is a solution in the healthcare domain, and our doctors would not be able to just rely on accuracy from a black box model.
- Rather, they would like to know more about how the predictions are being made, which would assist them in making better decisions.

Table 6.2 Heart disease dataset

Variable	Variable Type	Description
Age	Numeric	Patients Age in years
Sex	Nominal	Gender of patient <ul style="list-style-type: none">• Male – 1• Female – 0
Chest Pain Type	Nominal	Type of chest pain experienced by patient categorized into <ul style="list-style-type: none">• 1 – Typical• 2 – Typical angina• 3 – Non-anginal pain• 4 – Asymptomatic
resting bp s	Numerical	Level of blood pressure at resting mode in mm/HG
cholesterol	Numerical	Serum cholesterol in mg/dl
fasting blood sugar	Nominal	Blood sugar levels on fasting > 120 mg/dl represents as, <ul style="list-style-type: none">• 1 – True• 0 – false
resting ecg	Nominal	Result of electrocardiogram while at rest are represented in 3 distinct values <ul style="list-style-type: none">• 0 – Normal• 1 – Abnormality in ST-T wave• 2 – Left ventricular hypertrophy
max heart rate	Numerical	Maximum heart rate achieved
exercise angina	Nominal	Angina induced by exercise <ul style="list-style-type: none">• 0 – No• 1 – Yes
oldpeak	Numeric	Exercise induced ST-depression in comparison with the state of rest
ST slope	Nominal	ST segment measured in terms of slope during peak exercise <ul style="list-style-type: none">• 0 – Normal• 1 – Upsloping• 2 – Flat• 3 – Down sloping
target	Nominal	It is the target variable which we have to predict, <ul style="list-style-type: none">• 1 means patient is suffering from heart risk• 0 means patient is normal

Analytical Dashboard: Heart Disease Risk Prediction

- let us build an analytics dashboard which should aid:
 1. Data scientists to inspect the workings and performance of their model.
 2. Non-technical users to interactively inspect the inner workings of the model.
 3. Keep humans in the loop by ensuring we only accept the right decisions from the model and overrule the predictions in case of any missing information.
- Business analytics solutions such as Tableau provide us with the option of plugging our model outcome to a dashboard.
- For our example, we use the Plotly dash library built on Python to build our own custom analytics dashboard.
- Using ExplainerDashboard library built on Plotly dash [Dijk, 2019], we can create custom analytics dashboards with just a few lines of the python code.

Analytical Dashboard: Heart Disease Risk Prediction

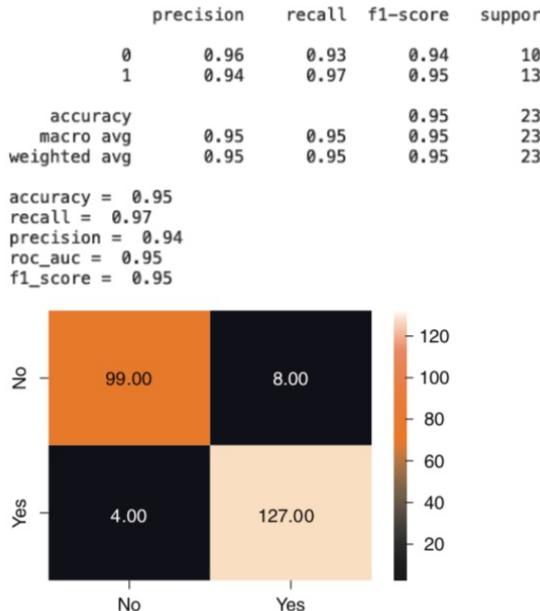


Figure 6.30 Random Forest model accuracy metrics.

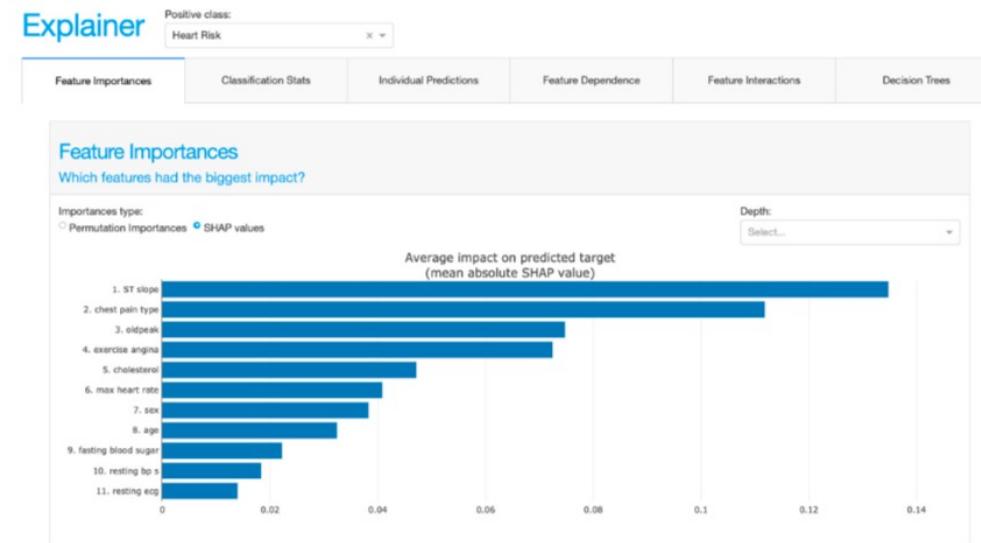
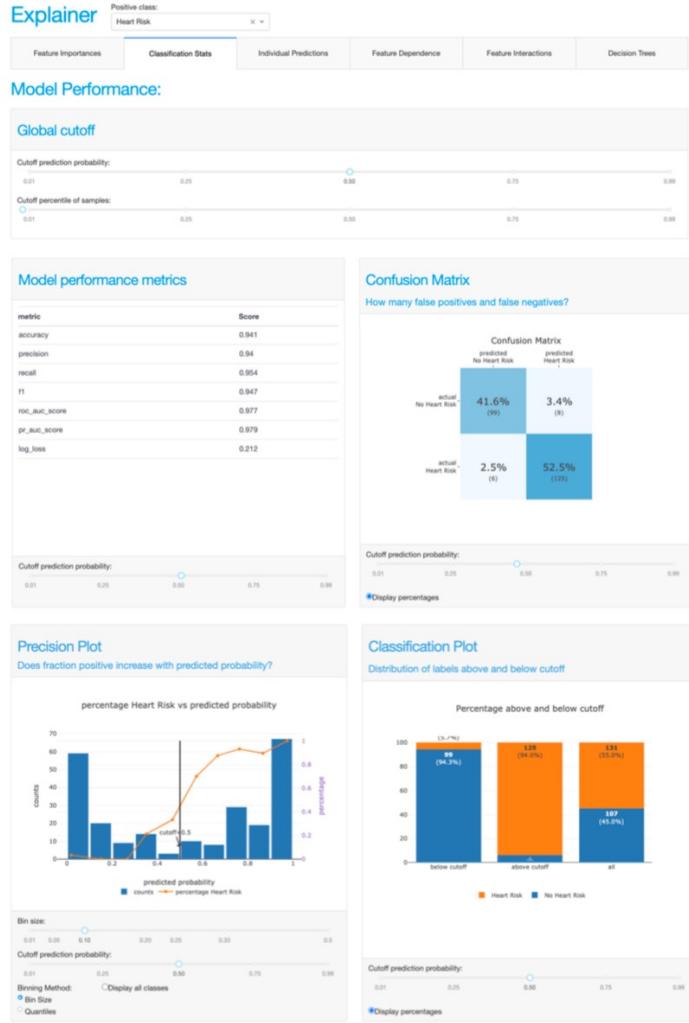


Figure 6.31 Analytics dashboard – feature importance.

- Refer to Fig. 6.30 for details on model metrics.
- Figure 6.31 explains the important features aiding the ML for predicting heart disease risk.

Analytical Dashboard: Heart Disease Risk Prediction



- Figures 6.32 and 6.33 explain model performance metrics.
- They also provide interactive filters and sliders for the user to check and adjust cut-off values.

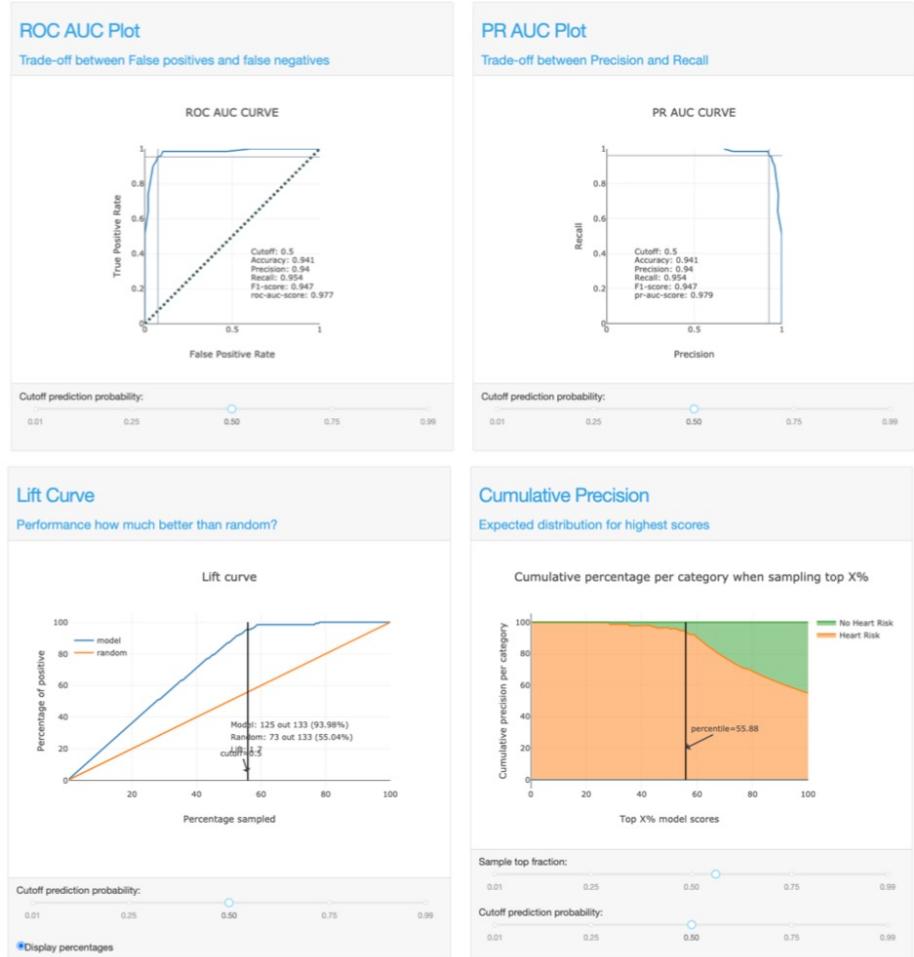


Figure 6.33 Analytics dashboard – model metrics.

Figure 6.32 Analytics dashboard – model metrics.

Analytical Dashboard: Heart Disease Risk Prediction

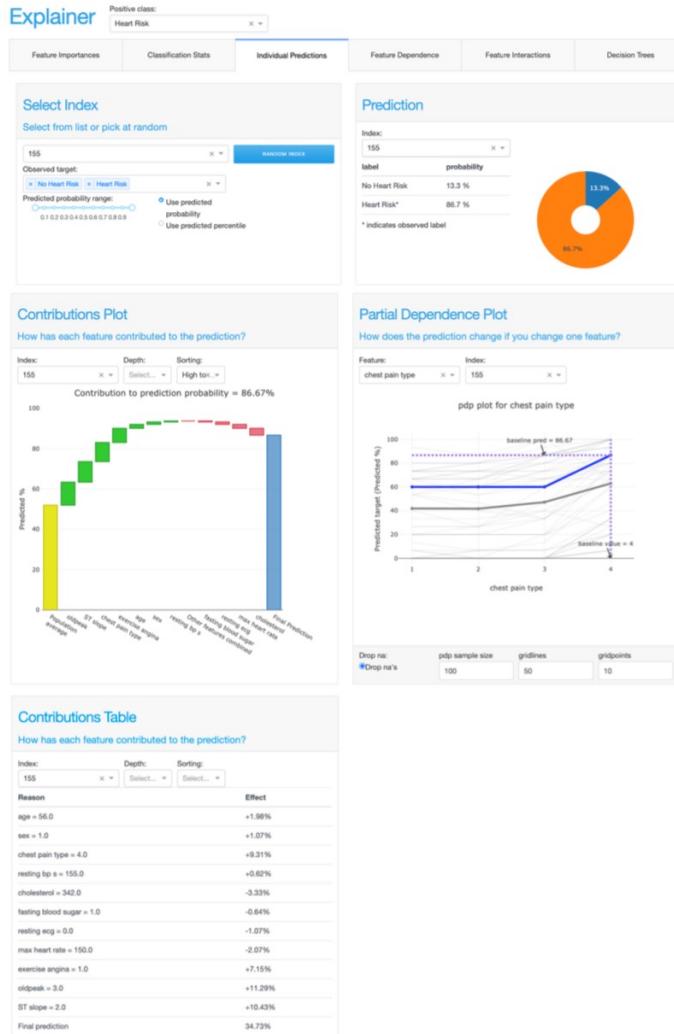


Figure 6.34 Analytics dashboard – individual predictions.

Refer Fig. 6.34)

- Using the Individual predictions tab, we can also dig deep into factors that helped predict a particular record.
- This would give the control back to humans which will enable them to take right decisions from the model.

Analytical Dashboard: Heart Disease Risk Prediction

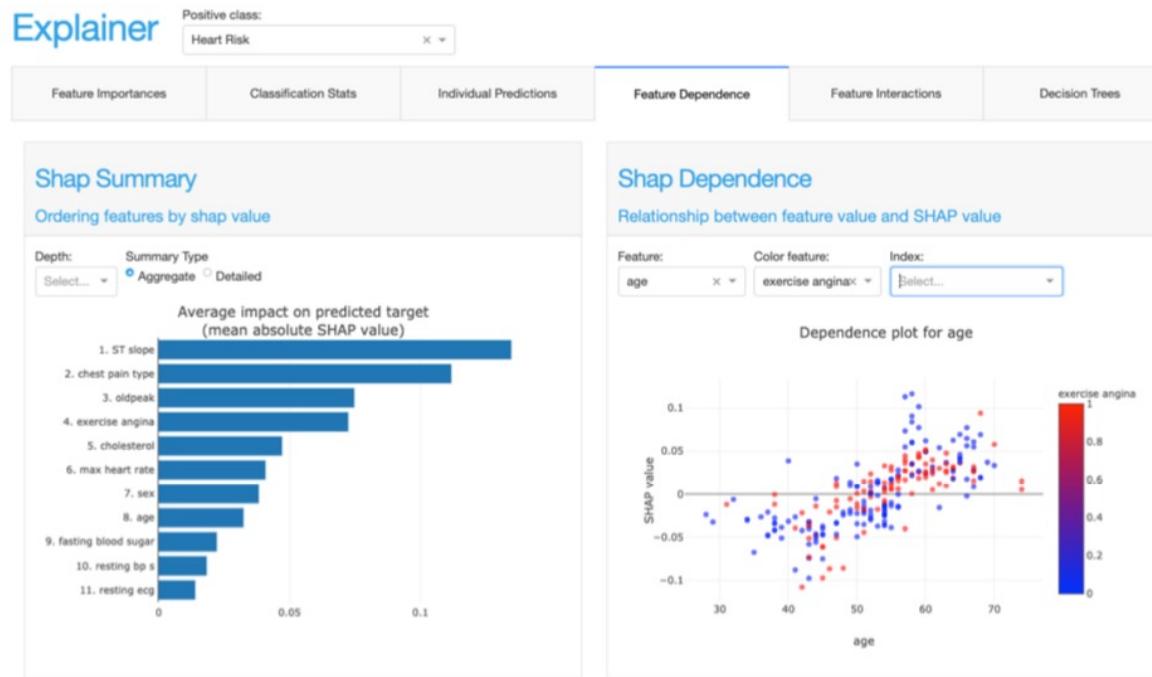
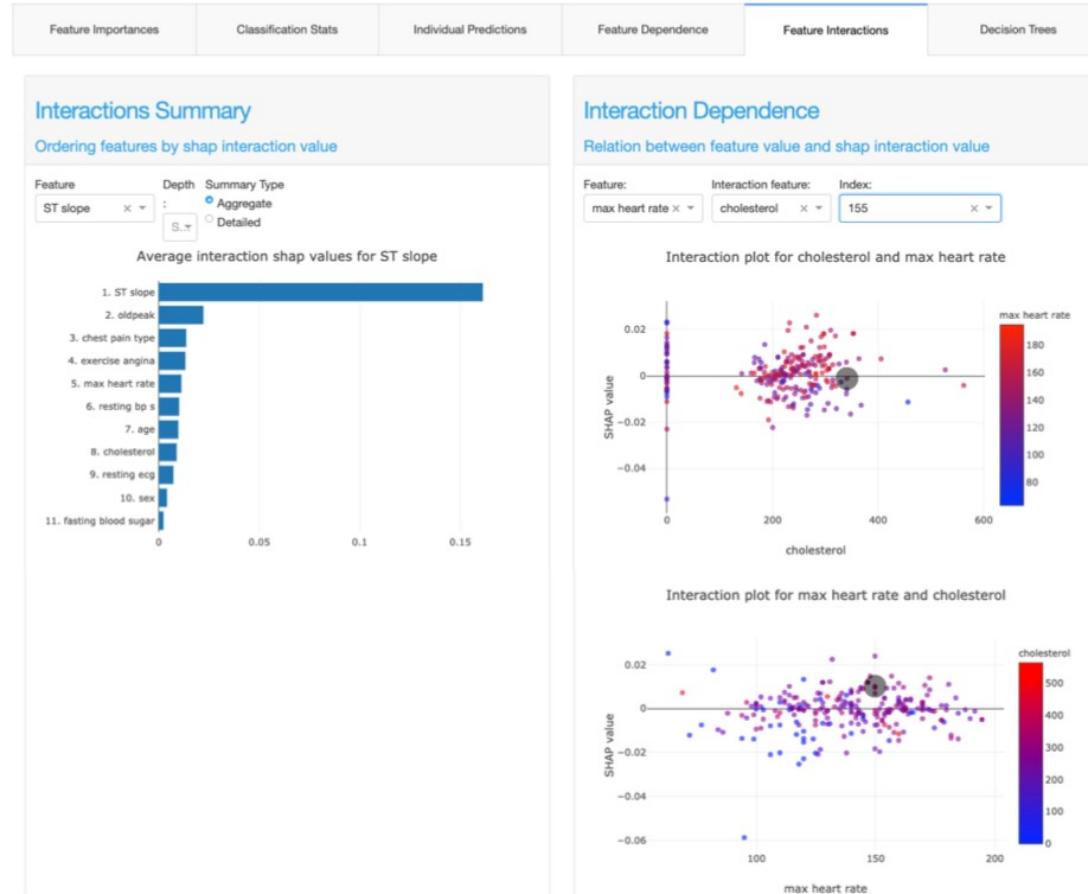


Figure 6.35 Analytics dashboard – feature dependencies.

- Refer to Fig. 6.35
- Feature dependencies tab helps us understand the relationship between multiple model features and SHAP explainer.
- SHAP value is used to illustrate how each features is affecting the model, as well as allows local and global analysis depending on the dataset and the problem.

Analytical Dashboard: Heart Disease Risk Prediction



- Refer to Fig. 6.36
- The feature Interactions tab helps us understand the relationship between multiple model features.
- By selecting an index value, we can also find where the particular patient record appears on the scatter plot.

Figure 6.36 Analytics dashboard – feature interactions.

Analytical Dashboard: Heart Disease Risk Prediction

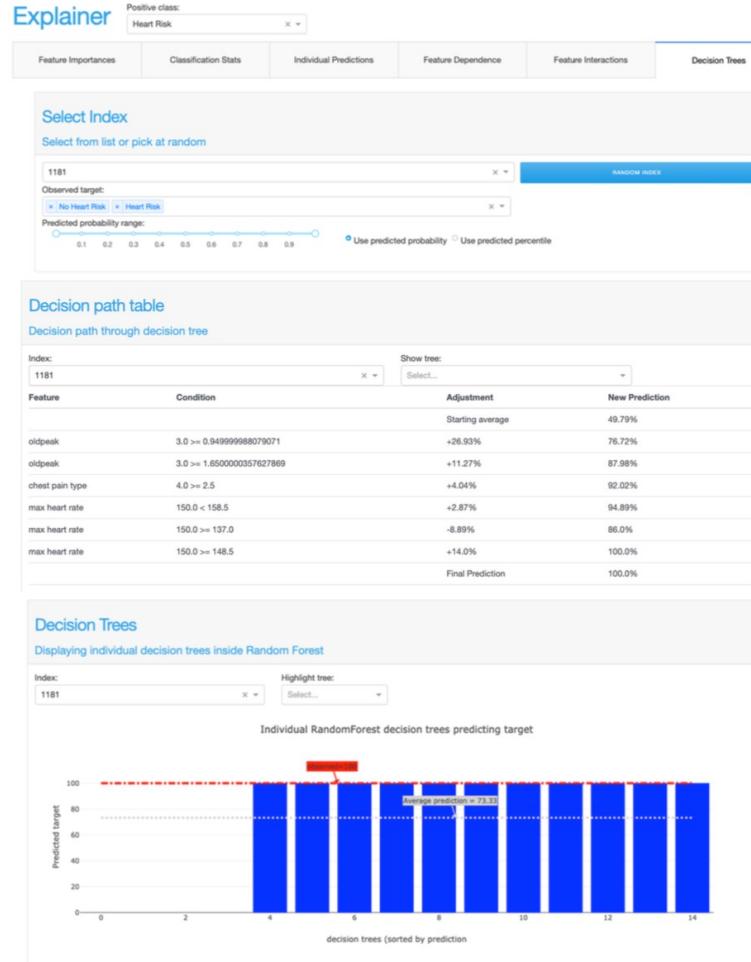


Figure 6.37 Analytics dashboard – decision trees.

- Refer to Figs. 6.37 and 6.38.
- They depict the **decision tree** through which the model predicts whether a particular record is at risk for heart disease.
- It even provides the complete decision tree structure and the path taken by the selected patient record.

Analytical Dashboard: Heart Disease Risk Prediction

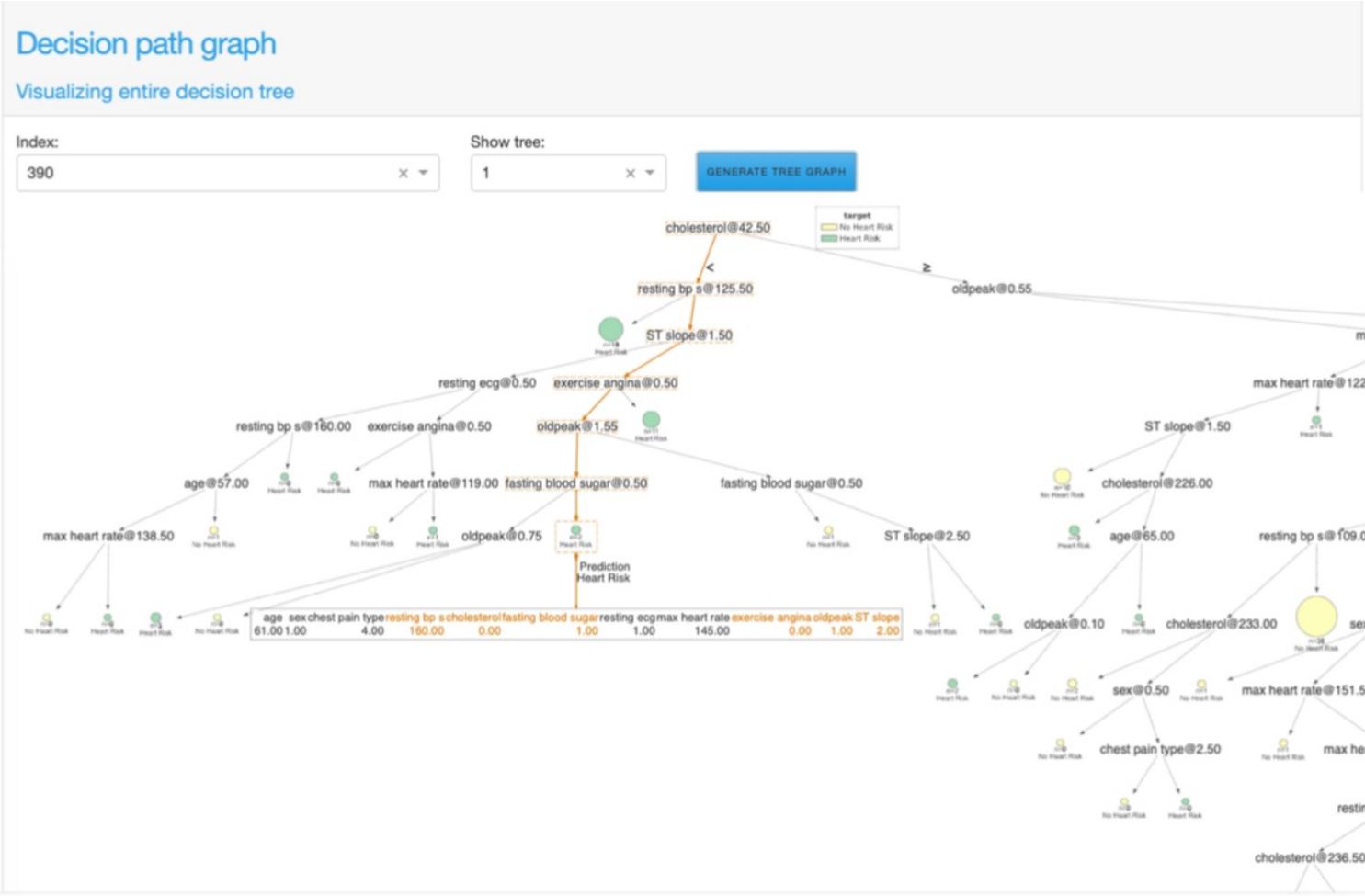


Figure 6.38 Analytics dashboard - decision trees.

Storytelling with Data – Key Elements

1. Keep it simple:

- When we tell stories during a presentation or while talking to a customer, in order to make our audience understand, we tend to use a lot of technical terms and business terminologies.
- This phenomenon of eventually talking in the language of the field is famously known as “the curse of knowledge”.
- The audience may not be aware of those technicalities about a project and start losing interest.
- Hence, we need to keep our story simple and clear to help our audience make sense of it.

Storytelling with Data – Key Elements

2. Highlight your message:

- Another important thing to incorporate in our presentation is to stay focused on the big idea we want to convey through our story.
- Our message should be the hero of our presentation.
- Each piece of information should lead our audience towards our big idea of the story.
- This should be the single most important information which we want our audience to remember.

Storytelling with Data – Key Elements

3. Clear and compact conclusion:

- When we can engage our audience in our presentation through a story, it becomes even more important to conclude it well.
- We may not always have a well-structured solution to our problem, but we should still be able to conclude our story.
 1. Either with the solution we have successfully implemented, or
 2. The learnings through our failures and the way forward accordingly.
- Leaving our audience hanging without a proper conclusion will never be able to serve the purpose of our presentation.
- Hence, it is necessary to conclude our story properly.

References:

- [Damasio, 2009] – Damasio (2009), A. Damasio, *When emotions make better decisions* available at https://www.youtube.com/watch?v=1wup_K2WN0I, last accessed August 15, 2020.
- [Snow, 1855] – Snow (1855), Snow, J., On the mode of communication of cholera (2nd ed.). London: John Churchill.
- [Stackpole, 2020] – Stackpole (2020), Beth Stackpole, The next chapter in analytics: *data storytelling*, available at <https://mitsloan.mit.edu/ideas-made-to-matter/next-chapter-analytics-data-story-telling>, May 2020.
- [Glenn, 2009] – J. Glenn (2009), J. Glenn and R. Walker, Significant objects and how they got that way, available at <https://significant-objects.com>, last accessed August 15, 2020.
- [Campbell, 2008] – J. Campbell (2008), Joseph Campbell, The Hero with a Thousand Faces (The collected works of Joseph Campbell).
- [Anon, 2020] – Anon (2020), The story mountain available at <https://www.agilify.co.uk/the-story-mountain/>, last accessed January 15, 2021.
- [Geljon, 2017] – M. Geljon (2017), M. Geljon, *How story structure can help you build better stories*, Marketing & growth hacking, available at <https://blog.markgrowth.com/how-story-structure-can-help-you-build-better-stories-cfc7f3993f7c>, last accessed September 20, 2020.
- [Cox, 2010] – Cox (2010), Amanda Cox, Budget Forecasts, Compared with Reality, Politics, The New York Times, available at <https://archive.nytimes.com/screenshots/www.nytimes.com/interactive/2010/02/02/us/politics/20100201-budget-porcupine-graphic.jpg>, last accessed September 25, 2020.
- [Ashkenas, 2014] – J. Ashkenas (2014), J. Ashkenas, can you live on the minimum Wage? Opinion, The New York Times, available at <https://www.nytimes.com/interactive/2014/02/09/opinion/minimum-wage.html>, last accessed September 30, 2020.
- [Jones, 2014] – B. Jones (2014), B. Jones, communicating data with Tableau: designing, developing, and delivering data visualizations, O'Reilly.
- [Janardhan, 2015] – A. Janardhan (2015), Arun Janardhan, Tamil Nadu's caste fields: In 1 year, over 100 murders in two southern districts, Web Address: <https://indianexpress.com/article/india/india-others/tns-caste-fields/>, last accessed December 28, 2021.
- [Kumar, 2021] – U. D. Kumar (2021), U. D. Kumar, Business analytics: *The science of data-driven decision making*, Wiley India.
- [Dijk, 2019] – O. Dijk (2019), Oege Dijk, explainer dashboard, available at <https://github.com/oegedijk/explainerdashboard>, last accessed February 5, 2021.
- [Segel, 2010] – Segel and Heer (2010), E. Segel and J. Heer, “Narrative Visualization: *Telling Stories with Data*”, in IEEE Transactions on Visualization and Computer Graphics, vol. 16, no. 6, pp. 1139–1148, November-December 2010, doi: 10.1109/TVCG.2010.179.

Thank You!