Data Science Process

Tushar B. Kute,
http://tusharkute.com

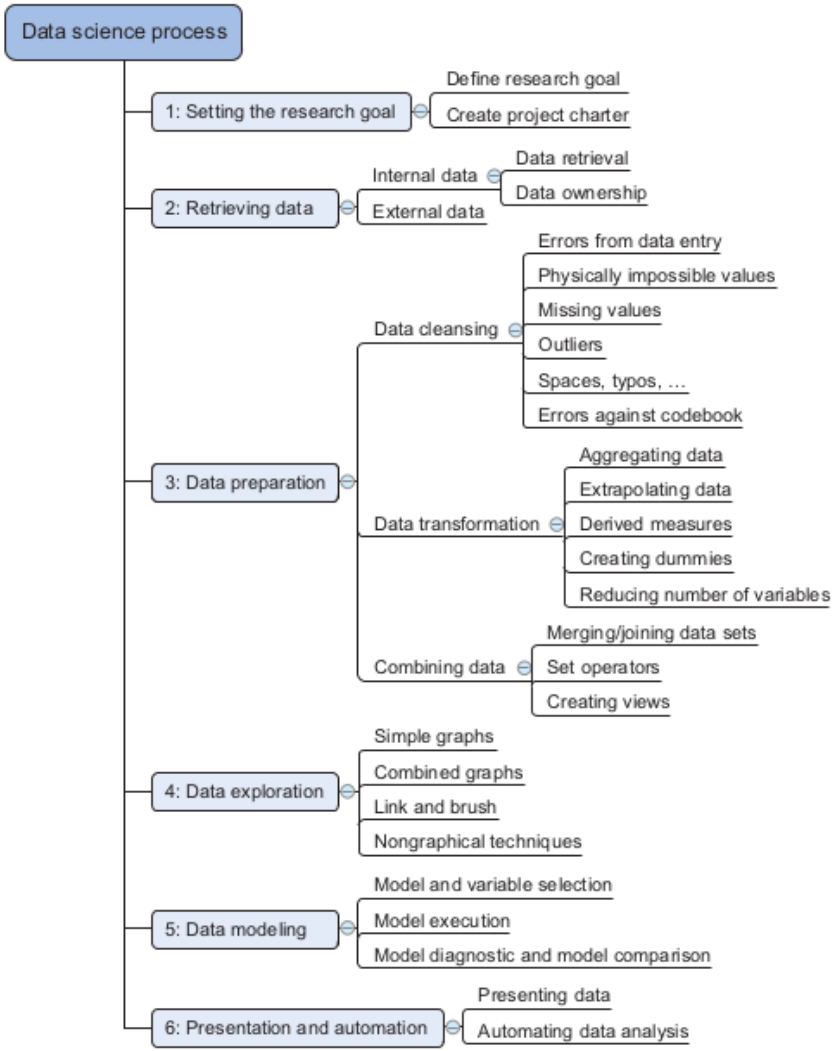
Objectives

- Understanding the flow of a data science process
- Discussing the steps in a data science process

Data Science Process

- A structured approach to data science helps you to maximize your chances of success in a data science project at the lowest cost.
- It also makes it possible to take up a project as a team, with each team member focusing on what they do best.
- However, this approach may not be suitable for every type of project or be the only way to do good data science.
- The typical data science process consists of six steps

Data Science Process



Steps

- The first step of this process is setting a research goal. The main purpose here is making sure all the stakeholders understand the what, how, and why of the project.
- The second phase is data retrieval. You want to have data available for analysis, so this step includes finding suitable data and getting access to the data from the data owner. The result is data in its raw form, which probably needs polishing and transformation before it becomes usable.

Steps

- Now that you have the raw data, it's time to prepare it. This includes transforming the data from a raw form into data that's directly usable in your models. To achieve this, you'll detect and correct different kinds of errors in the data, combine data from different data sources, and transform it. If you have successfully completed this step, you can progress to data visualization and modeling.
- The fourth step is data exploration. The goal of this step is to gain a deep understanding of the data. You'll look for patterns, correlations, and deviations based on visual and descriptive techniques. The insights you gain from this phase will enable you to start modeling.

Steps

- Finally, we get to the main part: model building (or “data modeling”).
- It is now that you attempt to gain the insights or make the predictions stated in your project charter.
- Now is the time to bring out the heavy guns, but remember research has taught us that often (but not always) a combination of simple models tends to outperform one complicated model.
- If you’ve done this phase right, you’re almost done.

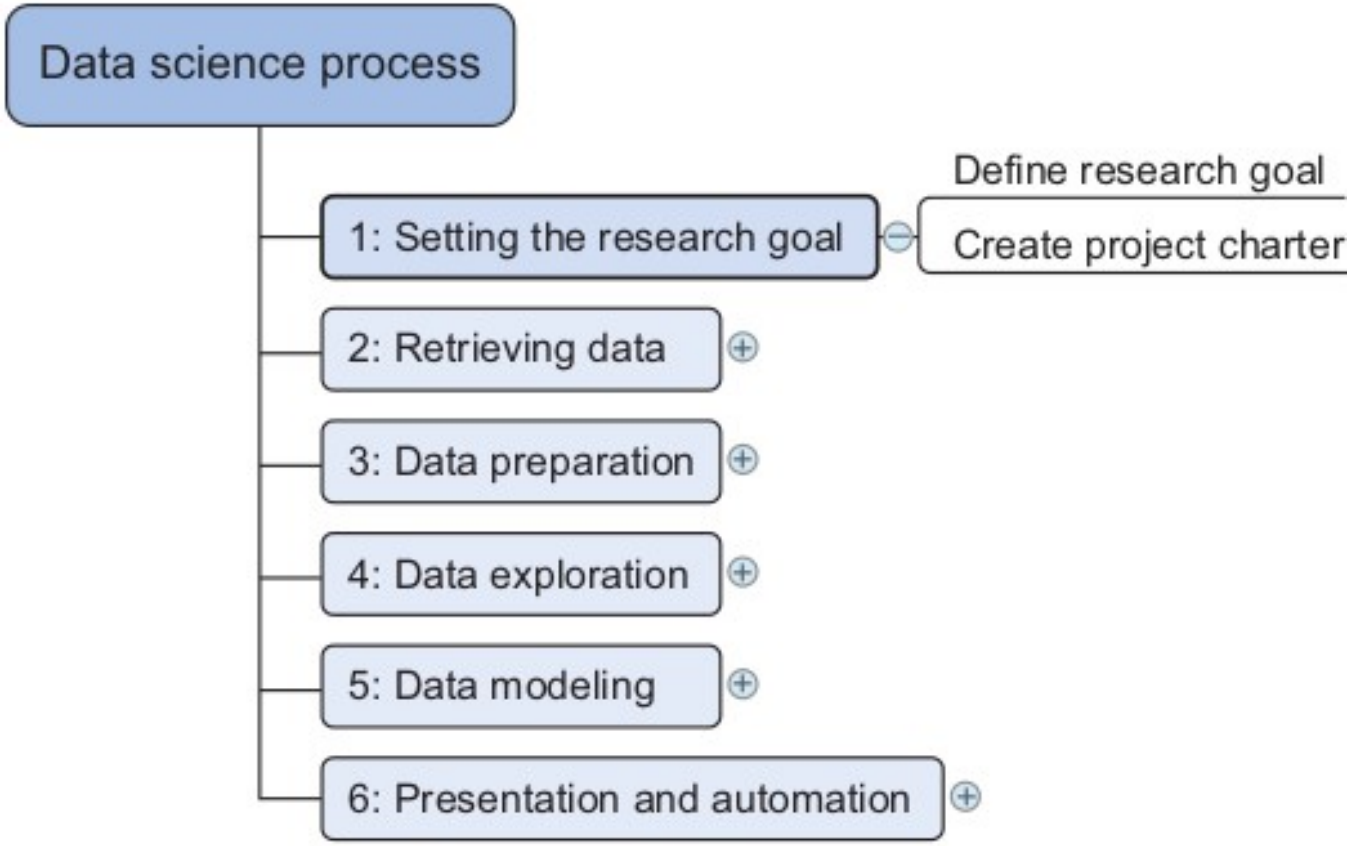
Steps

- The last step of the data science model is presenting your results and automating the analysis, if needed.
- One goal of a project is to change a process and/or make better decisions. You may still need to convince the business that your findings will indeed change the business process as expected.
- This is where you can shine in your influencer role. The importance of this step is more apparent in projects on a strategic and tactical level.
- Certain projects require you to perform the business process over and over again, so automating the project will save time.

Don’t be slave to this process

- Not every project will follow this blueprint, because your process is subject to the preferences of the data scientist, the company, and the nature of the project you work on.
- Some companies may require you to follow a strict protocol, whereas others have a more informal manner of working.
- In general, you’ll need a structured approach when you work on a complex project or when many people or resources are involved.

1. Setting research goal



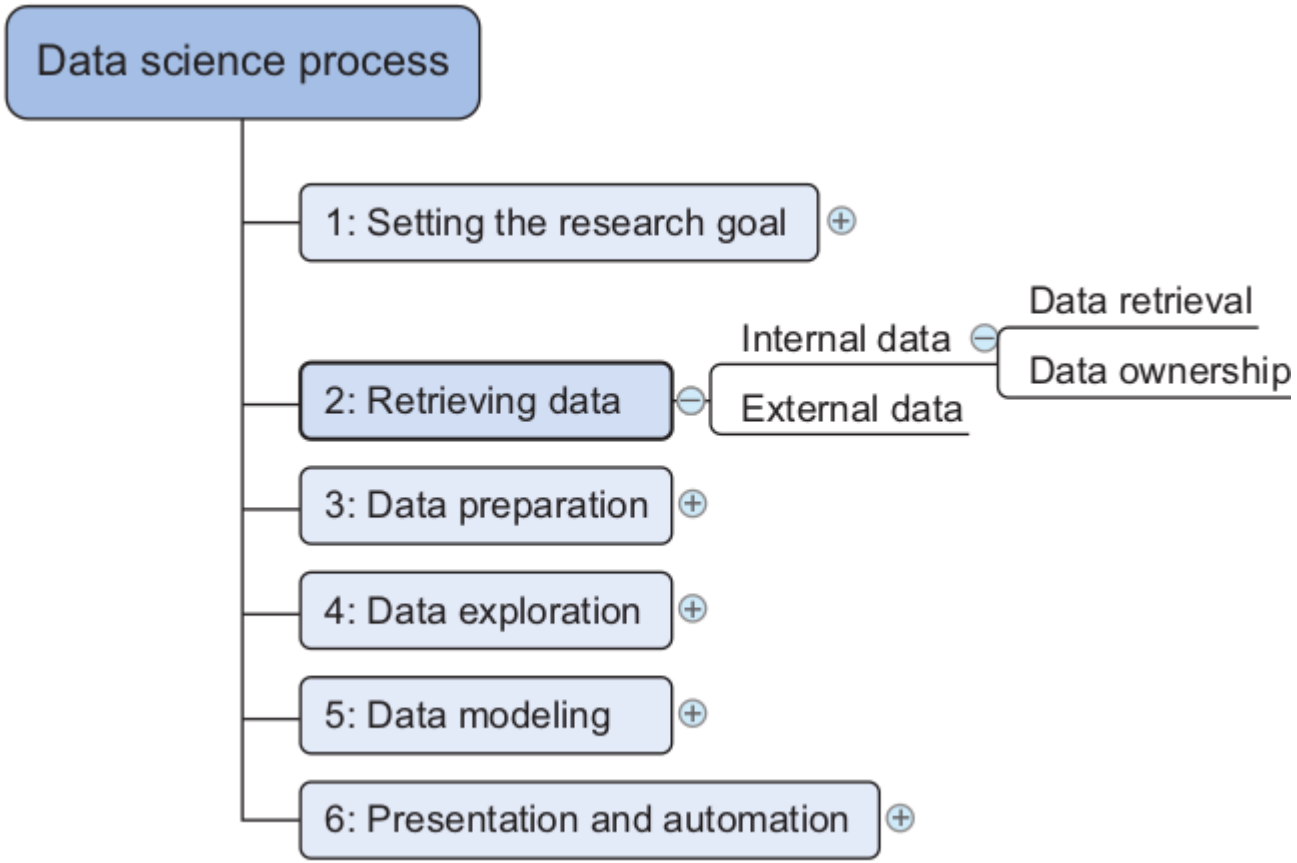
Goal and context of research

- An essential outcome is the research goal that states the purpose of your assignment in a clear and focused manner.
- Understanding the business goals and context is critical for project success.
- Continue asking questions and devising examples until you grasp the exact business expectations, identify how your project fits in the bigger picture, appreciate how your research is going to change the business, and understand how they'll use your results.

Create project charter

- Clients like to know upfront what they're paying for, so after you have a good understanding of the business problem, try to get a formal agreement on the deliverables. All this information is best collected in a project charter. For any significant project this would be mandatory.
- A project charter requires teamwork, and your input covers at least the following:
  - A clear research goal
  - The project mission and context
  - How you're going to perform your analysis
  - What resources you expect to use
  - Proof that it's an achievable project, or proof of concepts
  - Deliverables and a measure of success
  - A timeline

2. Retrieving data



Data Retrieval

- The next step in data science is to retrieve the required data. Sometimes you need to go into the field and design a data collection process yourself, but most of the time you won't be involved in this step.
- Many companies will have already collected and stored the data for you, and what they don't have can often be bought from third parties.
- Don't be afraid to look outside your organization for data, because more and more organizations are making even high-quality data freely available for public and commercial use.

Data Stored in company

- Your first act should be to assess the relevance and quality of the data that's readily available within your company.
- Most companies have a program for maintaining key data, so much of the cleaning work may already be done.
- This data can be stored in official data repositories such as databases, data marts, data warehouses, and data lakes maintained by a team of IT professionals.
- The primary goal of a database is data storage, while a data warehouse is designed for reading and analyzing that data.

Data Stored in company

- Getting access to data is another difficult task.
- Organizations understand the value and sensitivity of data and often have policies in place so everyone has access to what they need and nothing more.
- These policies translate into physical and digital barriers called Chinese walls. These “walls” are mandatory and well-regulated for customer data in most countries.
- This is for good reasons, too; imagine everybody in a credit card company having access to your spending habits.
- Getting access to the data may take time and involve company politics.

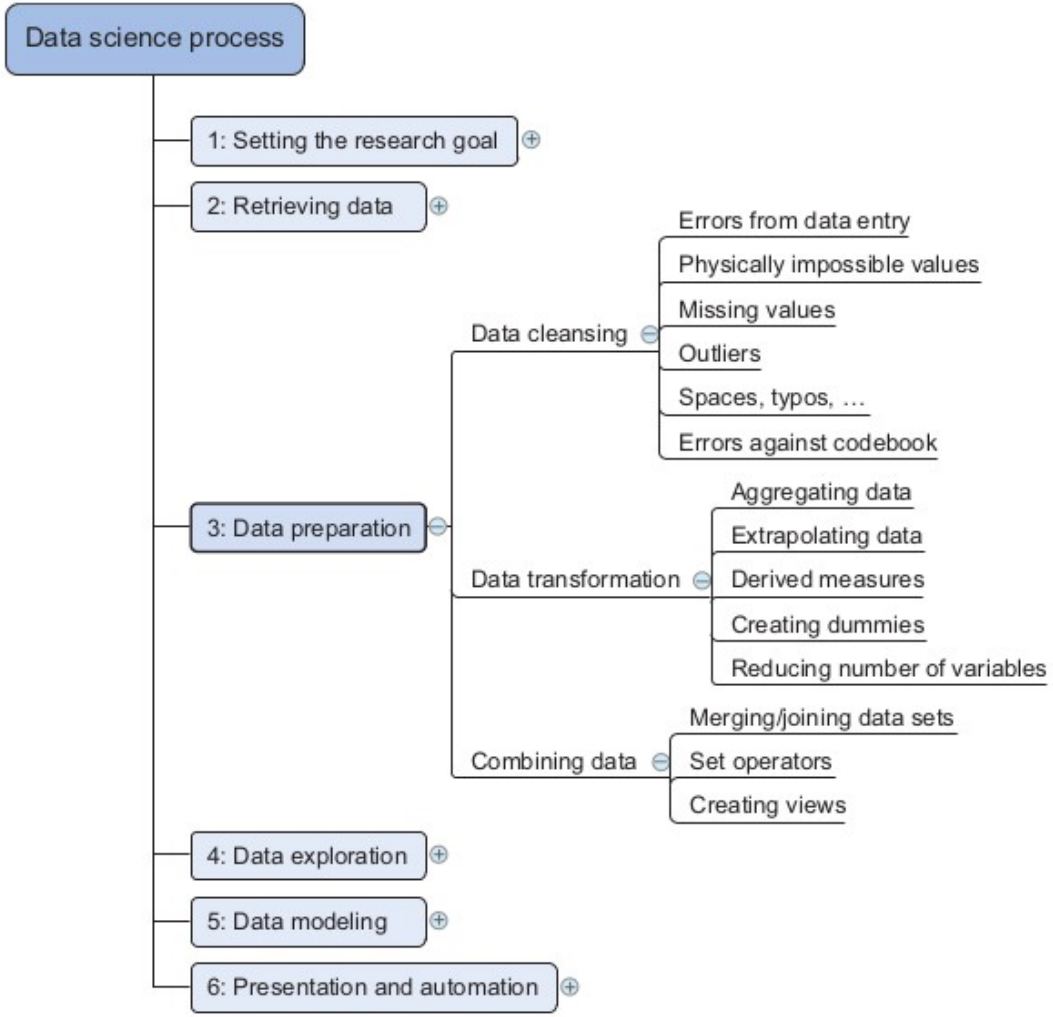
Data Sources

Open data site	Description
Data.gov	The home of the US Government's open data
https://open-data.europa.eu/	The home of the European Commission's open data
Freebase.org	An open database that retrieves its information from sites like Wikipedia, MusicBrains, and the SEC archive
Data.worldbank.org	Open data initiative from the World Bank
Aiddata.org	Open data for international development
Open.fda.gov	Open data from the US Food and Drug Administration

Data Quality Test

- Expect to spend a good portion of your project time doing data correction and cleansing, sometimes up to 80%.
- The retrieval of data is the first time you'll inspect the data in the data science process. Most of the errors you'll encounter during the data- gathering phase are easy to spot, but being too careless will make you spend many hours solving data issues that could have been prevented during data import.
- You'll investigate the data during the import, data preparation, and exploratory phases. The difference is in the goal and the depth of the investigation.

3. Data Preparation



Data Preparation

- The data received from the data retrieval phase is likely to be “a diamond in the rough.”
- Your task now is to sanitize and prepare it for use in the modeling and reporting phase.
- Doing so is tremendously important because your models will perform better and you’ll lose less time trying to fix strange output.
- It can’t be mentioned nearly enough times: garbage in equals garbage out.
- Your model needs the data in a specific format, so data transformation will always come into play.

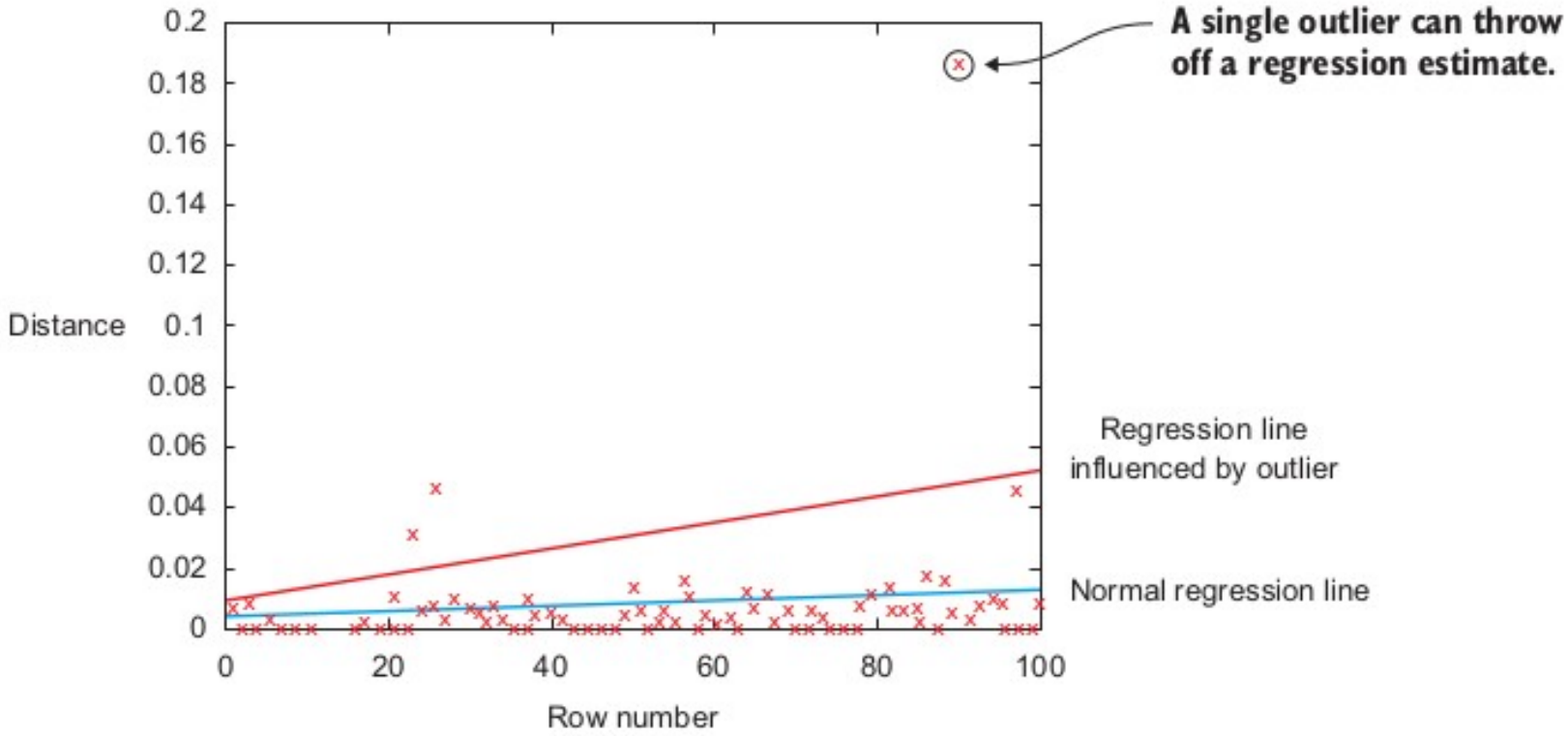
Data Cleansing

- Data cleansing is a subprocess of the data science process that focuses on removing errors in your data so your data becomes a true and consistent representation of the processes it originates from.
- By “true and consistent representation” we imply that at least two types of errors exist.
- The first type is the interpretation error, such as when you take the value in your data for granted, like saying that a person’s age is greater than 300 years.
- The second type of error points to inconsistencies between data sources or against your company’s standardized values.

Overview of common errors

General solution	
Try to fix the problem early in the data acquisition chain or else fix it in the program.	
Error description	**Possible solution**
Errors pointing to false values within one data set	
Mistakes during data entry	Manual overrules
Redundant white space	Use string functions
Impossible values	Manual overrules
Missing values	Remove observation or value
Outliers	Validate and, if erroneous, treat as missing value (remove or insert)
Errors pointing to inconsistencies between data sets	
Deviations from a code book	Match on keys or else use manual overrules
Different units of measurement	Recalculate
Different levels of aggregation	Bring to same level of measurement by aggregation or extrapolation

Example: Outliers



Data Entry Errors

- Data collection and data entry are error-prone processes.
- They often require human intervention, and because humans are only human, they make typos or lose their concentration for a second and introduce an error into the chain.
- But data collected by machines or computers isn't free from errors either. Errors can arise from human sloppiness, whereas others are due to machine or hardware failure.
- Examples of errors originating from machines are transmission errors or bugs in the extract, transform, and load phase (ETL).

Example: Frequency Table

Value	Count
Good	1598647
Bad	1354468
Godo	15
Bade	1

Error: Redundant Whitespaces

- Whitespaces tend to be hard to detect but cause errors like other redundant characters would.
- Who hasn’t lost a few days in a project because of a bug that was caused by whitespaces at the end of a string?
- You ask the program to join two keys and notice that observations are missing from the output file. After looking for days through the code, you finally find the bug.
- Then comes the hardest part: explaining the delay to the project stakeholders. The cleaning during the ETL phase wasn’t well executed, and keys in one table contained a whitespace at the end of a string.

Impossible values / Sanity Check

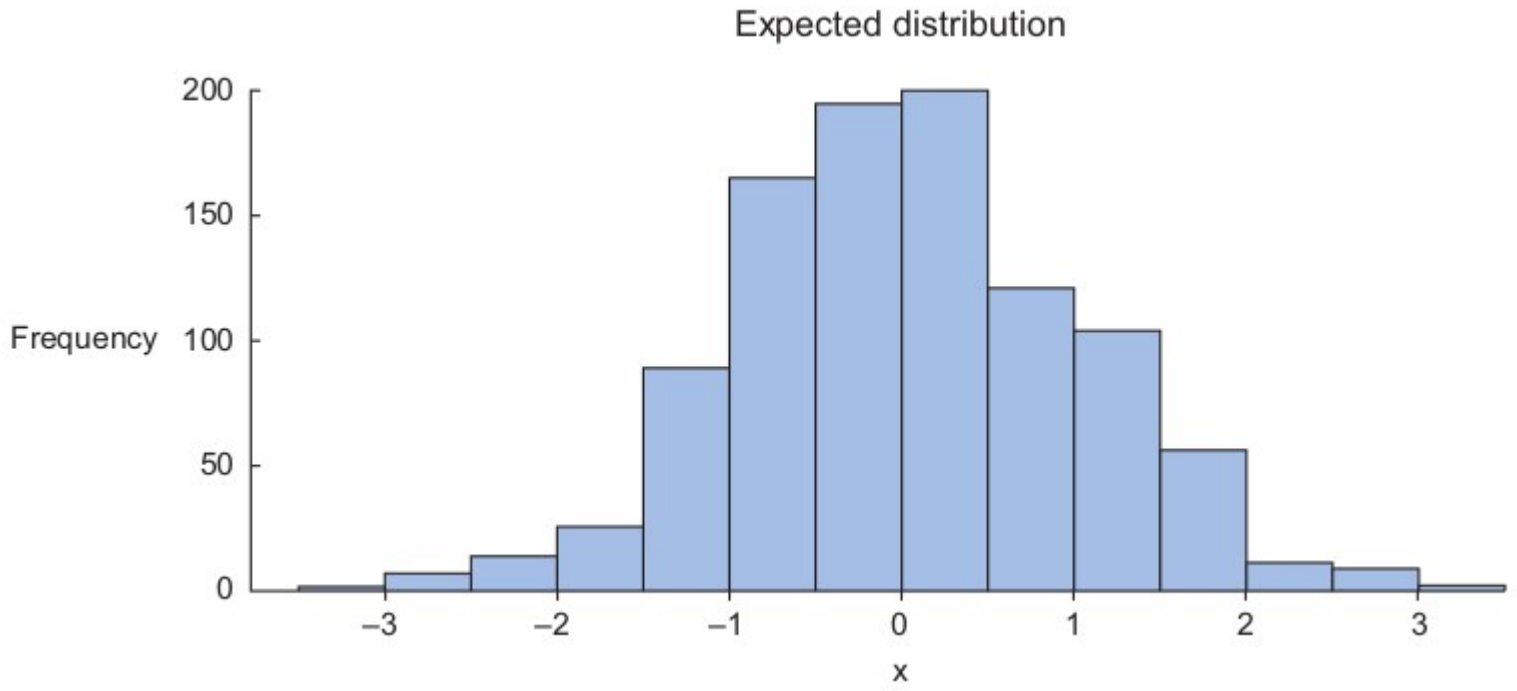
- Sanity checks are another valuable type of data check.
- Here you check the value against physically or theoretically impossible values such as people taller than 3 meters or someone with an age of 299 years.
- Sanity checks can be directly expressed with rules:

  check = 0 <= age <= 120

Outliers

- An outlier is an observation that seems to be distant from other observations or, more specifically, one observation that follows a different logic or generative process than the other observations.
- The easiest way to find outliers is to use a plot or a table with the minimum and maximum values.
- The plot on the top shows no outliers, whereas the plot on the bottom shows possible outliers on the upper side when a normal distribution is expected.
- The normal dis-tribution, or Gaussian distribution, is the most common distribution in natural sciences.

Example:



Example:


Distribution with outliers
Frequency
200
150
100
50
0
–2
0
2
4
y

Dealing with missing values

- Missing values aren’t necessarily wrong, but you still need to handle them separately; certain modeling techniques can’t handle missing values.
- They might be an indicator that something went wrong in your data collection or that an error happened in the ETL process.

Handling missing values

Technique	Advantage	Disadvantage
Omit the values	Easy to perform	You lose the information from an observation
Set value to `null`	Easy to perform	Not every modeling technique and/or implementation can handle `null` values
Impute a static value such as 0 or the mean	Easy to perform You don't lose information from the other variables in the observation	Can lead to false estimations from a model
Impute a value from an estimated or theoretical distribution	Does not disturb the model as much	Harder to execute You make data assumptions
Modeling the value (nondependent)	Does not disturb the model too much	Can lead to too much confidence in the model Can artificially raise dependence among the variables Harder to execute You make data assumptions

Error: deviation from code book

- Detecting errors in larger data sets against a code book or against standardized values can be done with the help of set operations.
- A code book is a description of your data, a form of metadata. It contains things such as the number of variables per observation, the number of observations, and what each encoding within a variable means.
- (For instance “0” equals “negative”, “5” stands for “very positive”.) A code book also tells the type of data you’re looking at: is it hierarchical, graph, something else?

Error: different units of measurement

- When integrating two data sets, you have to pay attention to their respective units of measurement.
- An example of this would be when you study the prices of gasoline in the world. To do this you gather data from different data providers.
- Data sets can con tain prices per gallon and others can contain prices per liter. A simple conversion will do the trick in this case.

Having different levels of aggregation

- Having different levels of aggregation is similar to having different types of measurement.
- An example of this would be a data set containing data per week versus one containing data per work week.
- This type of error is generally easy to detect, and summarizing (or the inverse, expanding) the data sets will fix it.
- After cleaning the data errors, you combine information from different data sources. But before we tackle this topic we'll take a little detour and stress the importance of cleaning data as early as possible.

Correct Errors

- A good practice is to mediate data errors as early as possible in the data collection chain and to fix as little as possible inside your program while fixing the origin of the problem.
- Retrieving data is a difficult task, and organizations spend millions of dollars on it in the hope of making better decisions.
- The data collection process is errorprone, and in a big organization it involves many steps and teams.

Correct Errors

- Data should be cleansed when acquired for many reasons:
  - Not everyone spots the data anomalies. Decision-makers may make costly mistakes on information based on incorrect data from applications that fail to correct for the faulty data.
  - If errors are not corrected early on in the process, the cleansing will have to be done for every project that uses that data.

Correct Errors

- Data errors may point to a business process that isn’t working as designed. For instance, both authors worked at a retailer in the past, and they designed a couponing system to attract more people and make a higher profit.
- Data errors may point to defective equipment, such as broken transmission lines and defective sensors.
- Data errors can point to bugs in software or in the integration of software that may be critical to the company.

Combine Data

- Your data comes from several different places, and in this substep we focus on integrating these different sources.
- Data varies in size, type, and structure, ranging from databases and Excel files to text documents.
- It’s easy to fill entire books on this topic alone, and we choose to focus on the data science process instead of presenting scenarios for every type of data.
- But keep in mind that other types of data sources exist, such as key-value stores, document stores, and so on, which we’ll handle in more appropriate places in the book.

Different ways to combine data

- You can perform two operations to combine information from different data sets.
- The first operation is joining: enriching an observation from one table with information from another table.
- The second operation is appending or stacking: adding the observations of one table to those of another table.
- When you combine data, you have the option to create a new physical table or a virtual table by creating a view. The advantage of a view is that it doesn’t consume more disk space.

Joining tables

- Joining tables allows you to combine the information of one observation found in one table with the information that you find in another table. The focus is on enriching a single observation.
- Let’s say that the first table contains information about the purchases of a customer and the other table contains information about the region where your customer lives.
- Joining the tables allows you to combine the information so that you can use it for your model

Joining tables

Client	Item	Month
John Doe	Coca-Cola	January
Jackie Qi	Pepsi-Cola	January

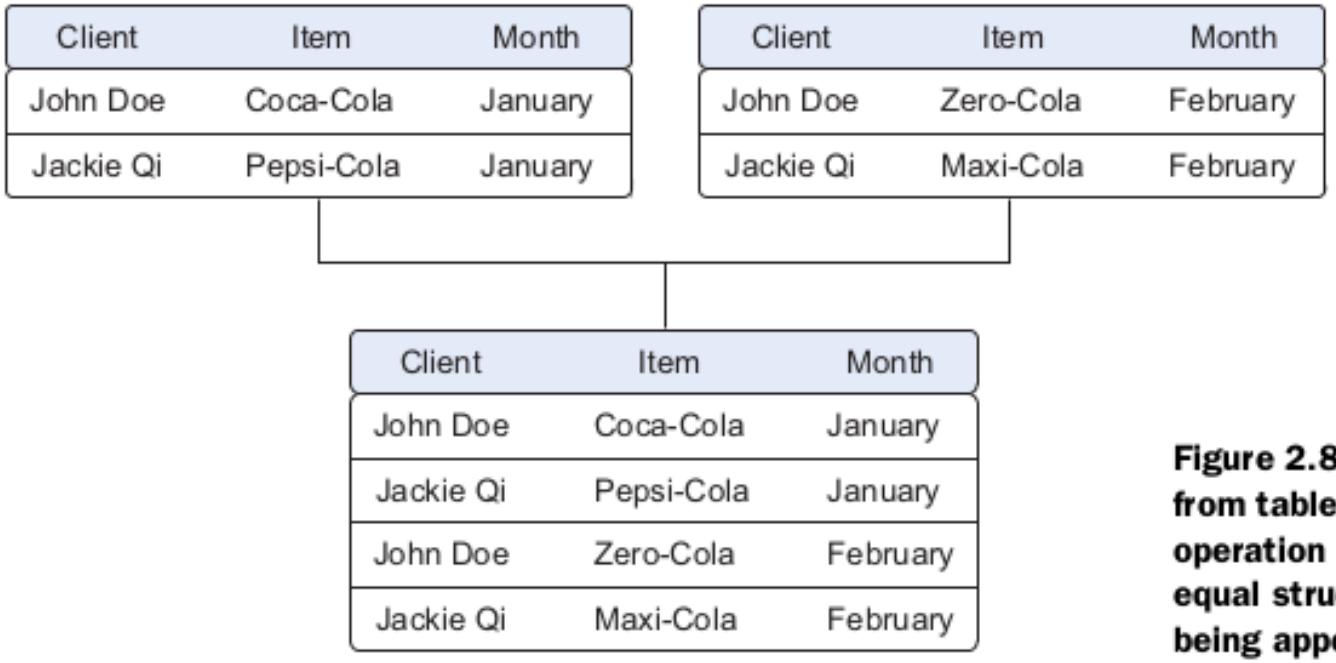
Client	Region
John Doe	NY
Jackie Qi	NC

Client	Item	Month	Region
John Doe	Coca-Cola	January	NY
Jackie Qi	Pepsi-Cola	January	NC

Appending tables

- Appending or stacking tables is effectively adding observations from one table to another table. One table contains the observations from the month January and the second table contains observations from the month February.
- The result of appending these tables is a larger one with the observations from January as well as February.
- The equivalent operation in set theory would be the union, and this is also the command in SQL, the common language of relational databases.
- Other set operators are also used in data science, such as set difference and intersection.

Appending tables



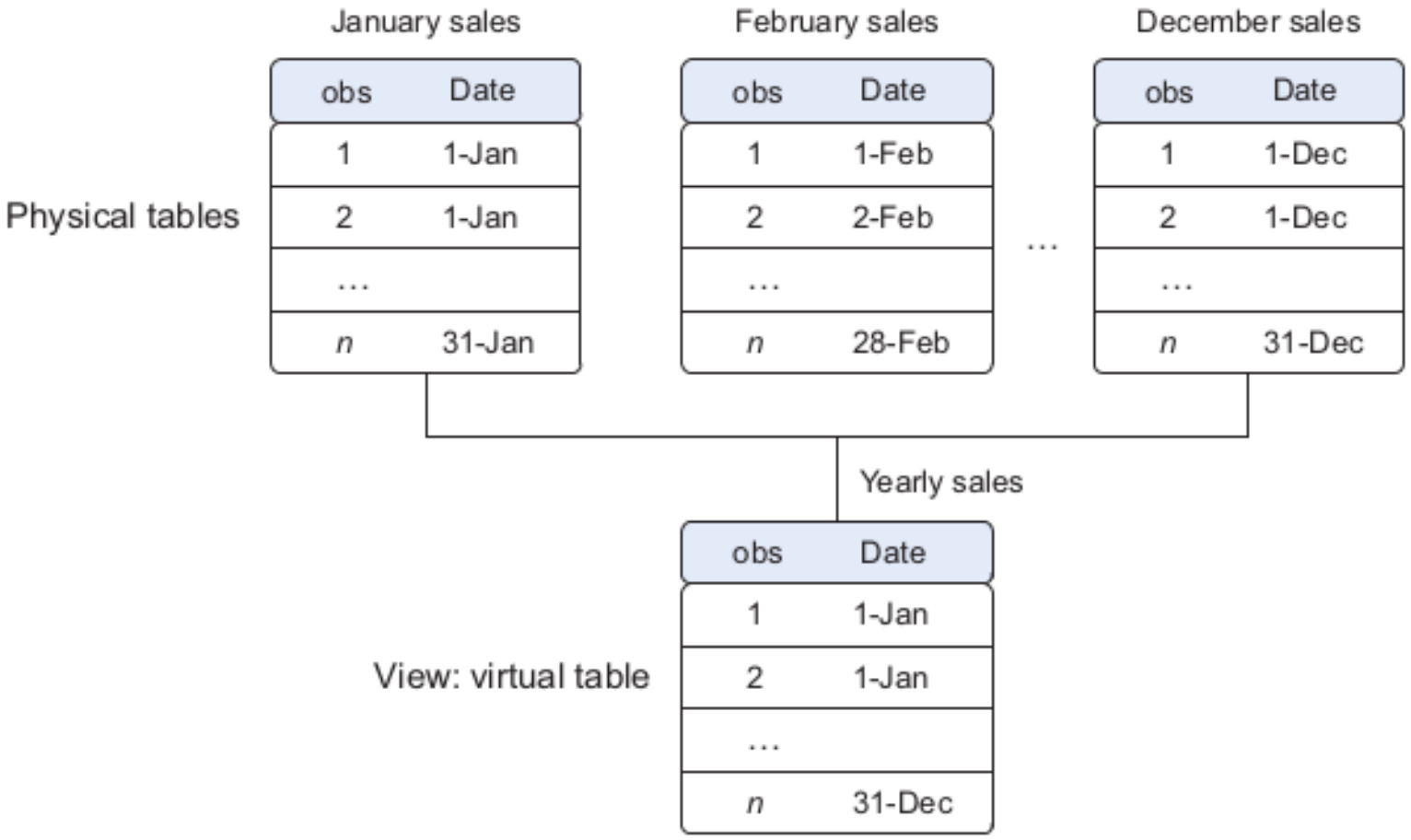
Client	Item	Month
John Doe	Coca-Cola	January
Jackie Qi	Pepsi-Cola	January

Client	Item	Month
John Doe	Zero-Cola	February
Jackie Qi	Maxi-Cola	February

Client	Item	Month
John Doe	Coca-Cola	January
Jackie Qi	Pepsi-Cola	January
John Doe	Zero-Cola	February
Jackie Qi	Maxi-Cola	February

Figure 2.8 Appending data from tables is a common operation but requires an equal structure in the tables being appended.

View: without replication



Aggregating measures

- Data enrichment can also be done by adding calculated information to the table, such as the total number of sales or what percentage of total stock has been sold in a certain region.
- Extra measures such as these can add perspective. Looking at figure, we now have an aggregated data set, which in turn can be used to calculate the participation of each product within its category.
- This could be useful during data exploration but more so when creating data models.

Example:

Product class	Product	Sales in $	Sales t-1 in $	Growth	Sales by product class	Rank sales
A	B	X	Y	(X-Y) / Y	AX	NX
Sport	Sport 1	95	98	–3.06%	215	2
Sport	Sport 2	120	132	–9.09%	215	1
Shoes	Shoes 1	10	6	66.67%	10	3

Data Transformation

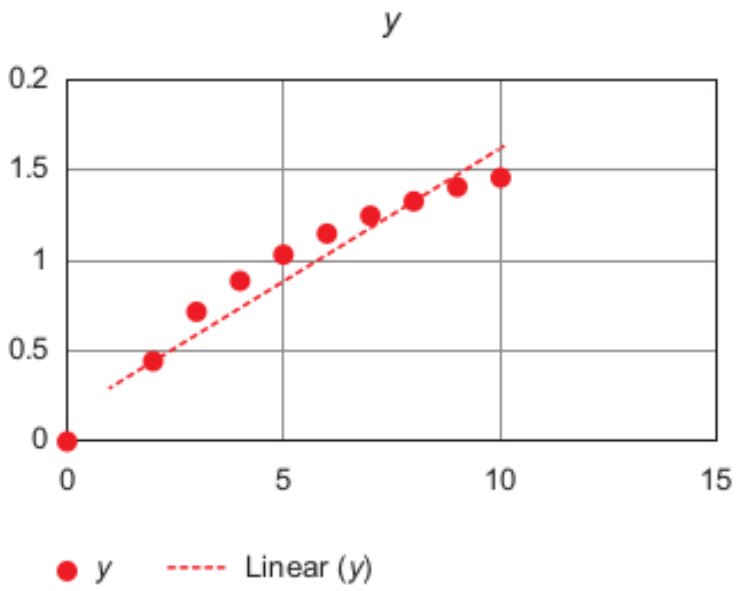
- Certain models require their data to be in a certain shape.
- Now that you've cleansed and integrated the data, this is the next task you'll perform: transforming your data so it takes a suitable form for data modeling.

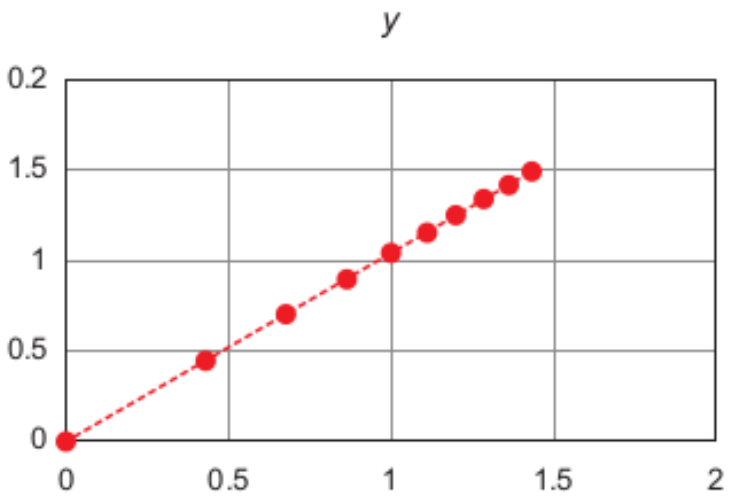
Data Transformation

- Relationships between an input variable and an output variable aren't always linear.
- Take, for instance, a relationship of the form $y = ae^{bx}$.
- Taking the log of the independent variables simplifies the estimation problem dramatically.

Data Transformation

x	1	2	3	4	5	6	7	8	9	10
log(x)	0.00	0.43	0.68	0.86	1.00	1.11	1.21	1.29	1.37	1.43
y	0.00	0.44	0.69	0.87	1.02	1.11	1.24	1.32	1.38	1.46





Reducing number of variables

- Sometimes you have too many variables and need to reduce the number because they don’t add new information to the model.
- Having too many variables in your model makes the model difficult to handle, and certain techniques don’t perform well when you overload them with too many input variables.

Reducing number of variables

- For instance, all the techniques based on a Euclidean distance perform well only up to 10 variables.

Euclidean distance

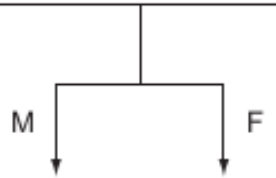
Euclidean distance or "ordinary" distance is an extension to one of the first things anyone learns in mathematics about triangles (trigonometry): Pythagoras's leg theorem. If you know the length of the two sides next to the 90° angle of a right-angled triangle you can easily derive the length of the remaining side (hypotenuse). The formula for this is hypotenuse = $\sqrt{(side1 + side2)^2}$. The Euclidean distance between two points in a two-dimensional plane is calculated using a similar formula: distance = $\sqrt{((x1 - x2)^2 + (y1 - y2)^2)}$. If you want to expand this distance calculation to more dimensions, add the coordinates of the point within those higher dimensions to the formula. For three dimensions we get distance = $\sqrt{((x1 - x2)^2 + (y1 - y2)^2 + (z1 - z2)^2)}$.

Dummy Variables

- Variables can be turned into dummy variables (figure). Dummy variables can only take two values: true(1) or false(0).
- They’re used to indicate the absence of a categorical effect that may explain the observation. In this case you’ll make separate columns for the classes stored in one variable and indicate it with 1 if the class is present and 0 otherwise.
- An example is turning one column named Weekdays into the columns Monday through Sunday. You use an indicator to show if the observation was on a Monday; you put 1 on Monday and 0 elsewhere.
- Turning variables into dummies is a technique that’s used in modeling and is popular with, but not exclusive to, economists.

Dummy Variables

Customer	Year	Gender	Sales
1	2015	F	10
2	2015	M	8
1	2016	F	11
3	2016	M	12
4	2017	F	14
3	2017	M	13

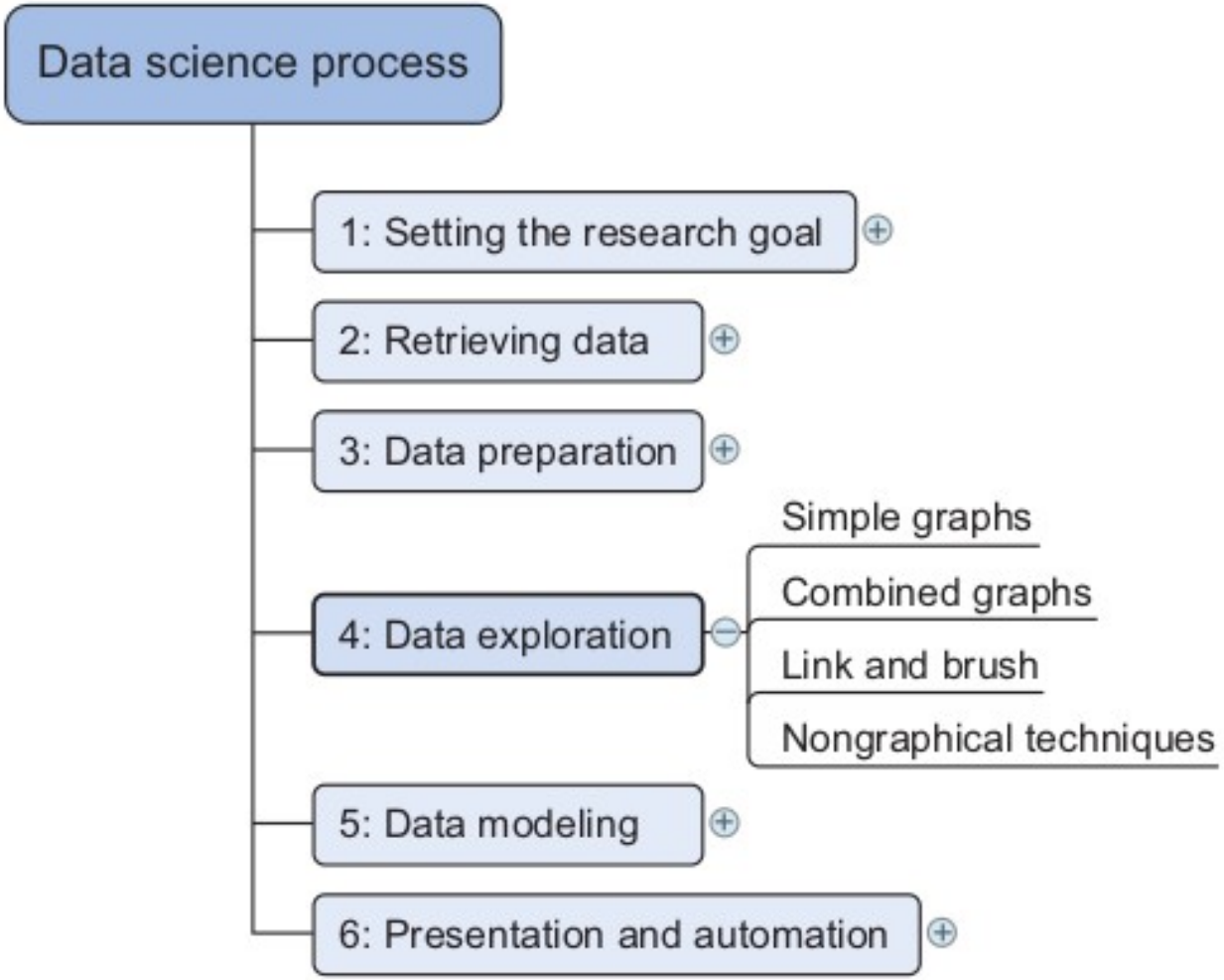


Customer	Year	Sales	Male	Female
1	2015	10	0	1
1	2016	11	0	1
2	2015	8	1	0
3	2016	12	1	0
3	2017	13	1	0
4	2017	14	0	1

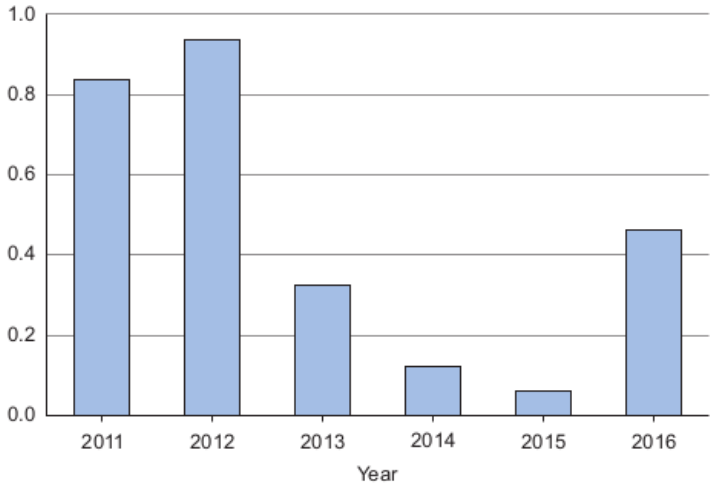
4. Exploratory Data Analysis

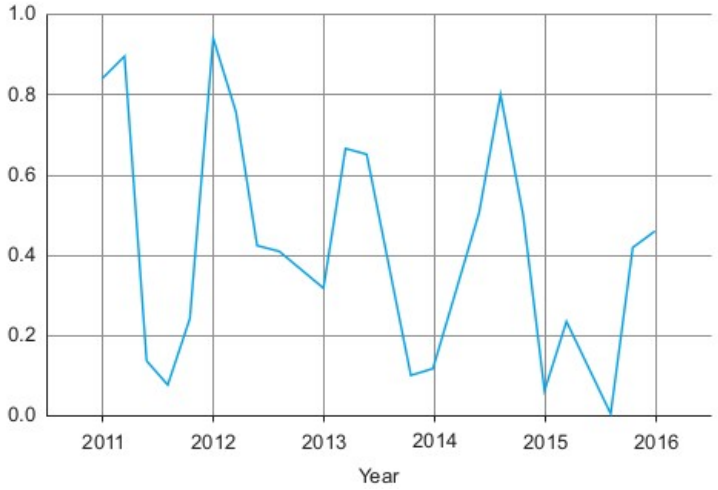
- During exploratory data analysis you take a deep dive into the data (see figure).
- Information becomes much easier to grasp when shown in a picture, therefore you mainly use graphical techniques to gain an understanding of your data and the interactions between variables.
- This phase is about exploring data, so keeping your mind open and your eyes peeled is essential during the exploratory data analysis phase.
- The goal isn't to cleanse the data, but it's common that you'll still discover anomalies you missed before, forcing you to take a step back and fix them.

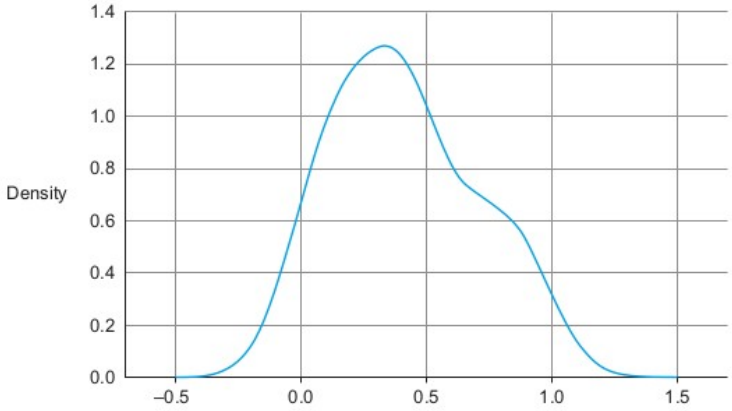
4. Exploratory Data Analysis



Exploratory Data Analysis



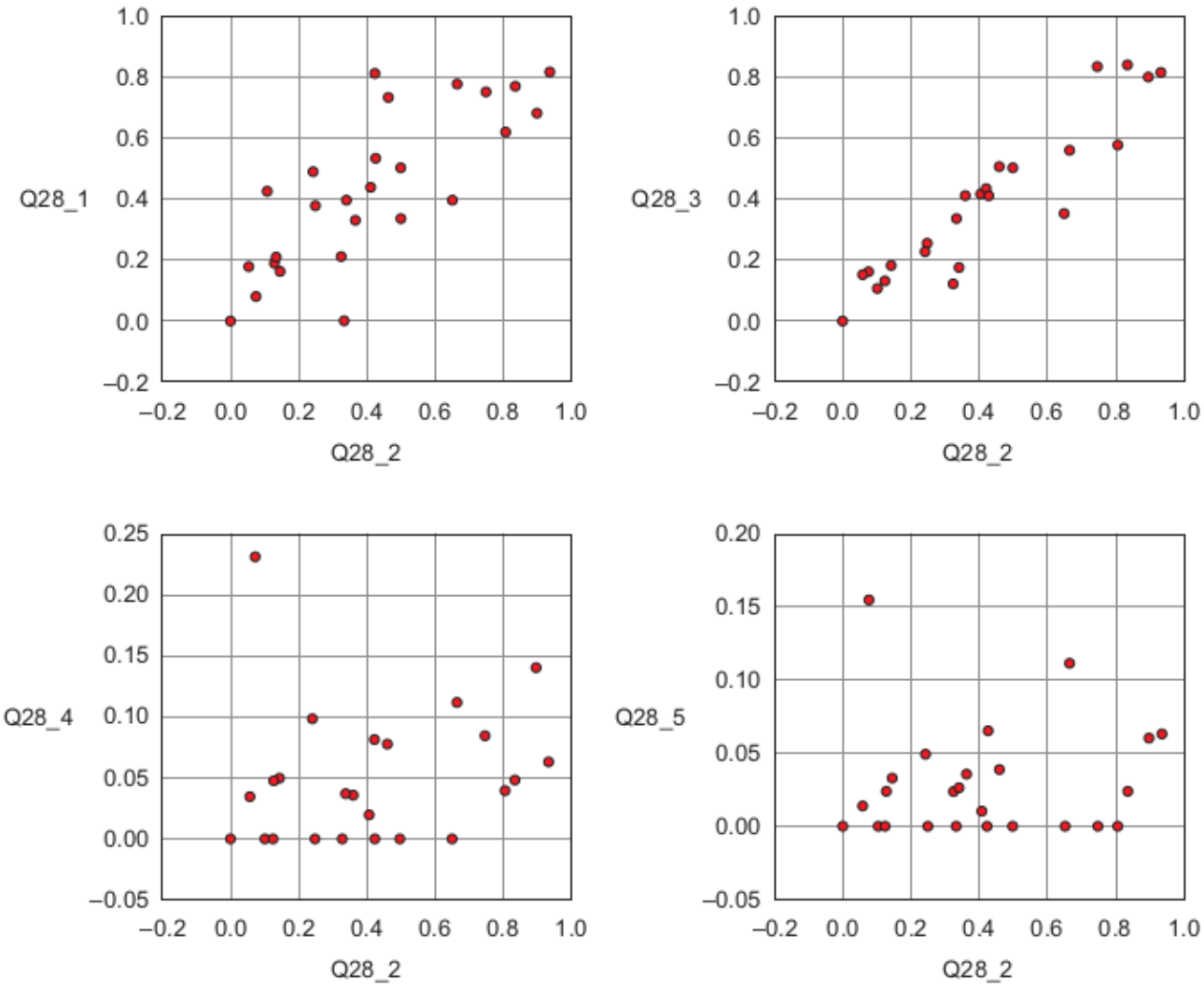




Brushing and linking

- With brushing and linking you combine and link different graphs and tables (or views) so changes in one graph are automatically transferred to the other graphs.

Brushing and linking



Brushing and linking

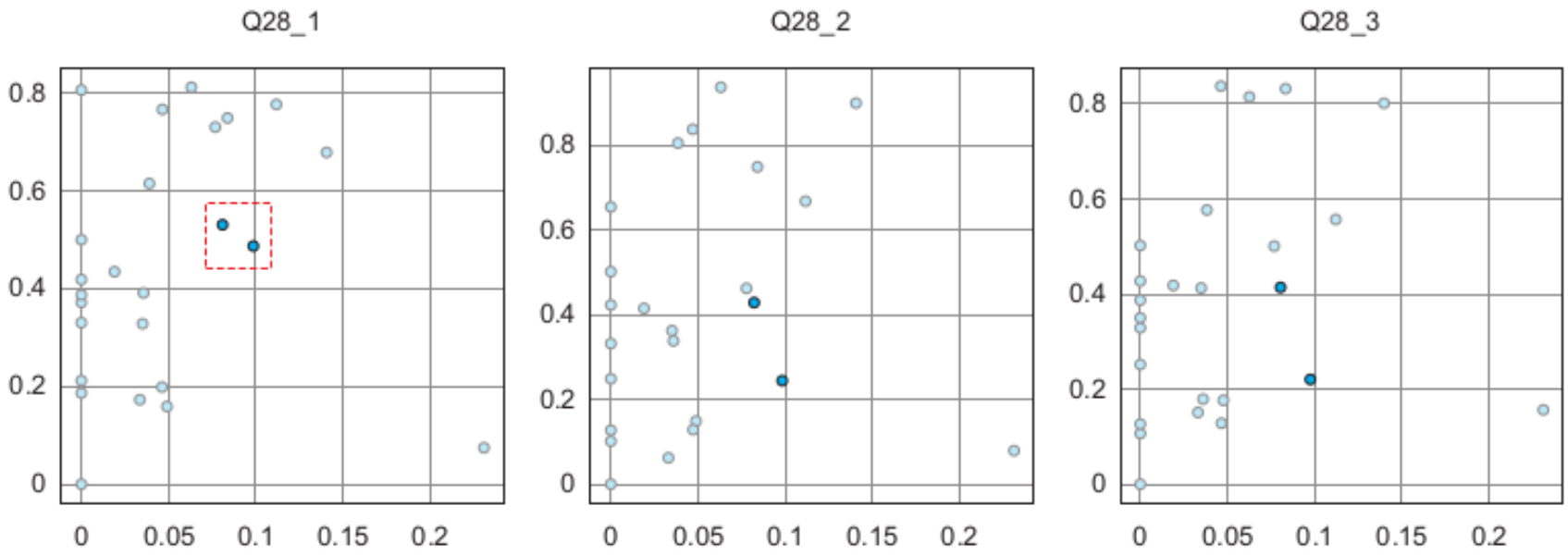
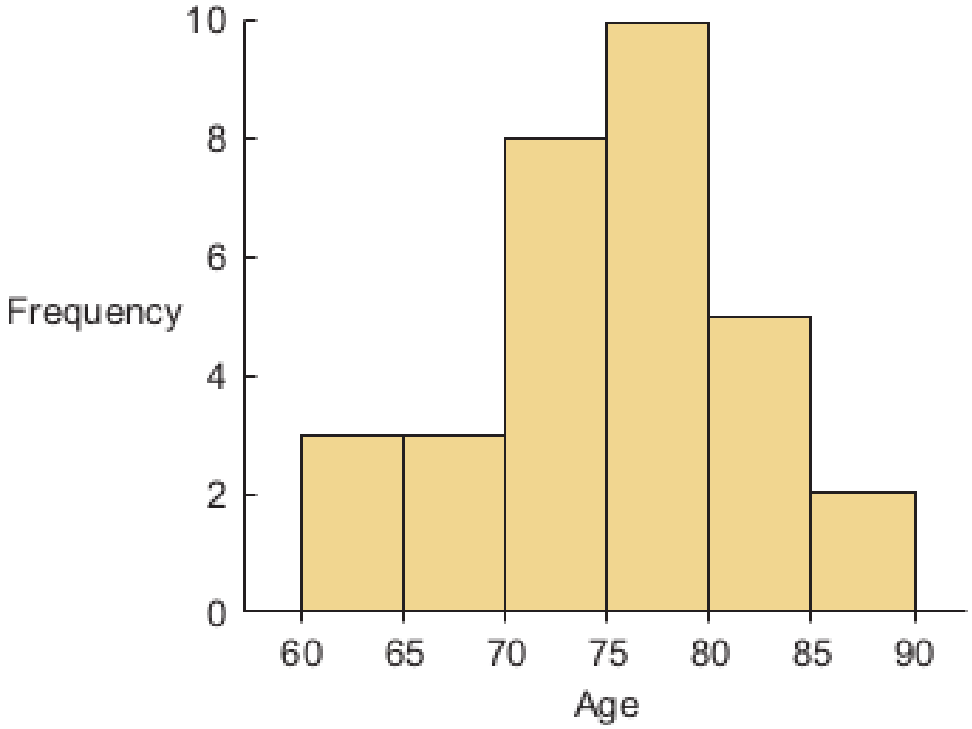


Figure 2.18 Link and brush allows you to select observations in one plot and highlight the same observations in the other plots.

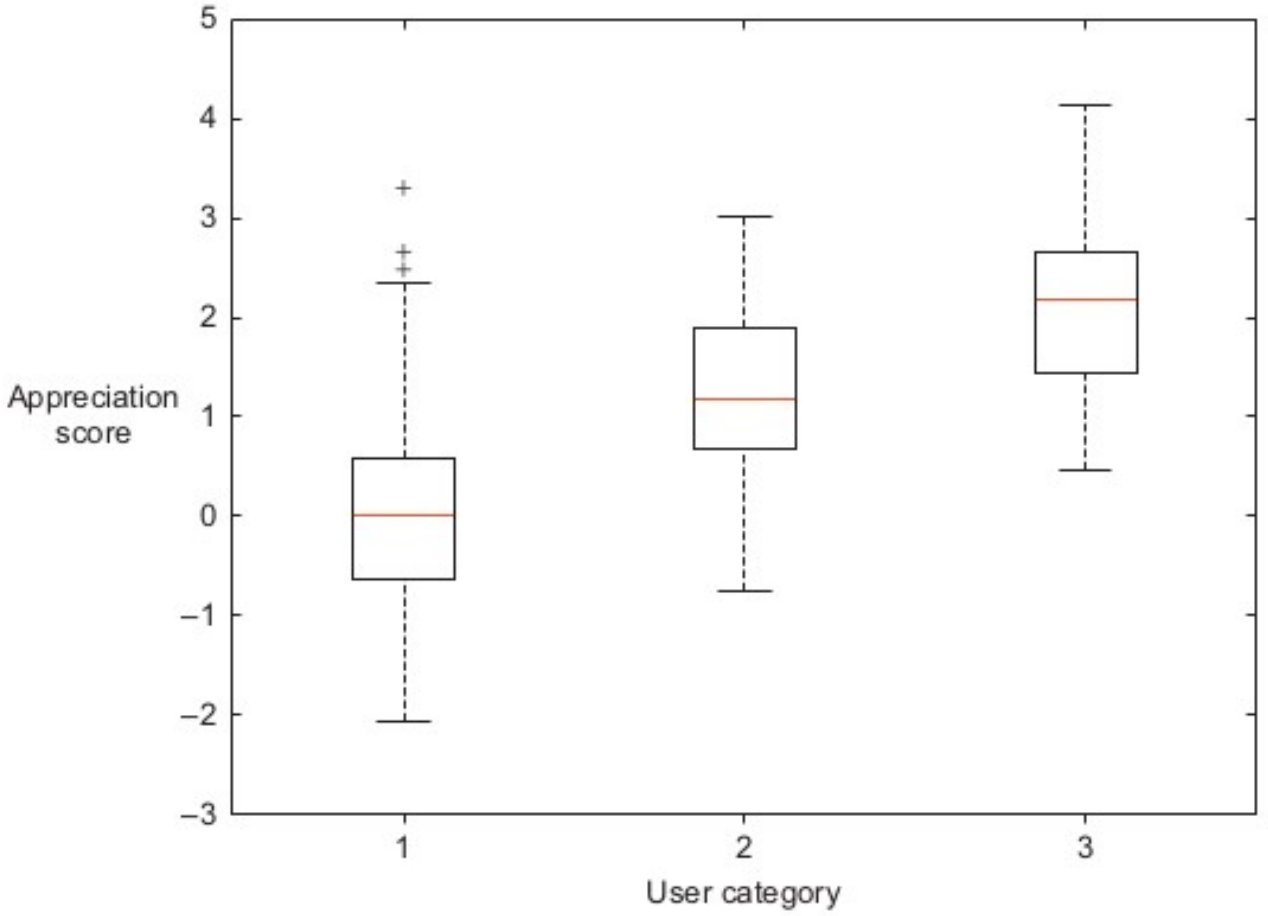
Histogram

- In a histogram a variable is cut into discrete categories and the number of occurrences in each category are summed up and shown in the graph.
- The boxplot, on the other hand, doesn’t show how many observations are present but does offer an impression of the distribution within categories.
- It can show the maximum, minimum, median, and other characterizing measures at the same time.

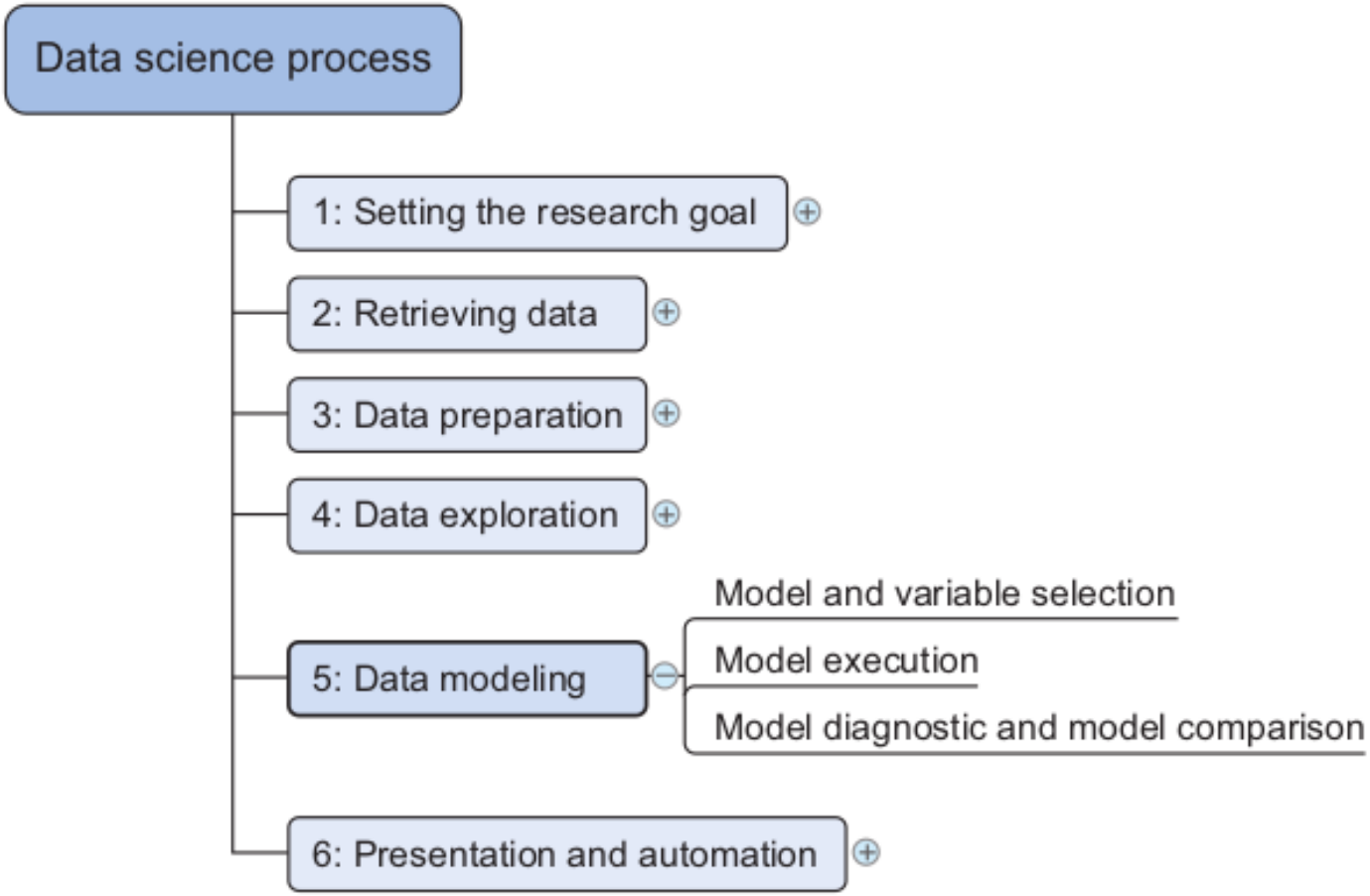
Histogram



Boxplot



5. Build the model



Building a model

- With clean data in place and a good understanding of the content, you're ready to build models with the goal of making better predictions, classifying objects, or gaining an understanding of the system that you're modeling.
- This phase is much more focused than the exploratory analysis step, because you know what you're looking for and what you want the outcome to be.

Building a model

- Building a model is an iterative process. The way you build your model depends on whether you go with classic statistics or the somewhat more recent machine learning school, and the type of technique you want to use.
- Either way, most models consist of the following main steps:
  - Selection of a modeling technique and variables to enter in the model
  - Execution of the model
  - Diagnosis and model comparison

Build a model

- You'll need to select the variables you want to include in your model and a modeling technique.
- Your findings from the exploratory analysis should already give a fair idea of what variables will help you construct a good model.
- Many modeling techniques are available, and choosing the right model for a problem requires judgment on yourpart.

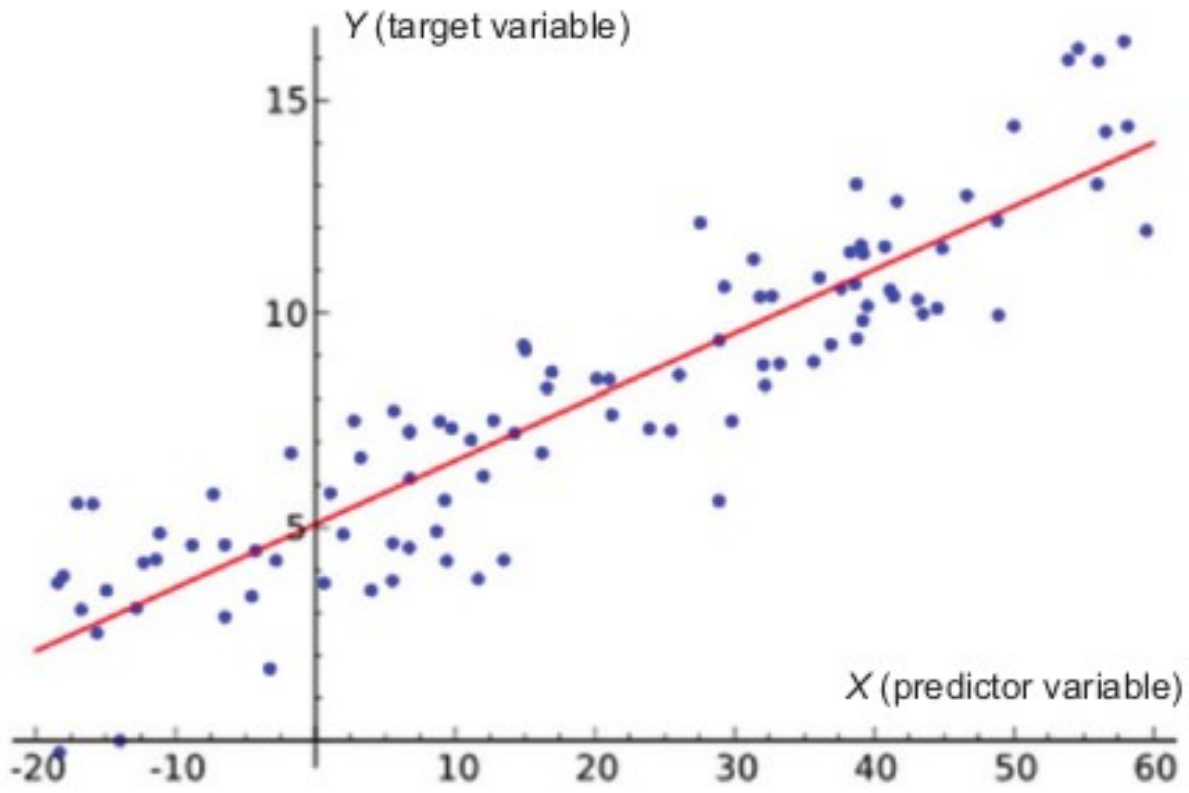
Build a model

- You'll need to consider model performance and whether your project meets all the requirements to use your model, as well as other factors:
  - Must the model be moved to a production environment and, if so, would it be easy to implement?
  - How difficult is the maintenance on the model: how long will it remain relevant if left untouched?
  - Does the model need to be easy to explain?

Model Execution

- Luckily, most programming languages, such as Python, already have libraries such as StatsModels or Scikit-learn. These packages use several of the most popular techniques.
- Coding a model is a nontrivial task in most cases, so having these libraries available can speed up the process.
- As you can see in the following code, it’s fairly easy to use linear regression (figure) with StatsModels or Scikit-learn.
- Doing this your self would require much more effort even for the simple techniques.

Model Execution



Coding

Listing 2.1 Executing a linear prediction model on semi-random data

```
import statsmodels.api as sm
import numpy as np
predictors = np.random.random(1000).reshape(500,2)
target = predictors.dot(np.array([0.4, 0.6])) + np.random.random(500)
lmRegModel = sm.OLS(target,predictors)
result = lmRegModel.fit()
result.summary()
```

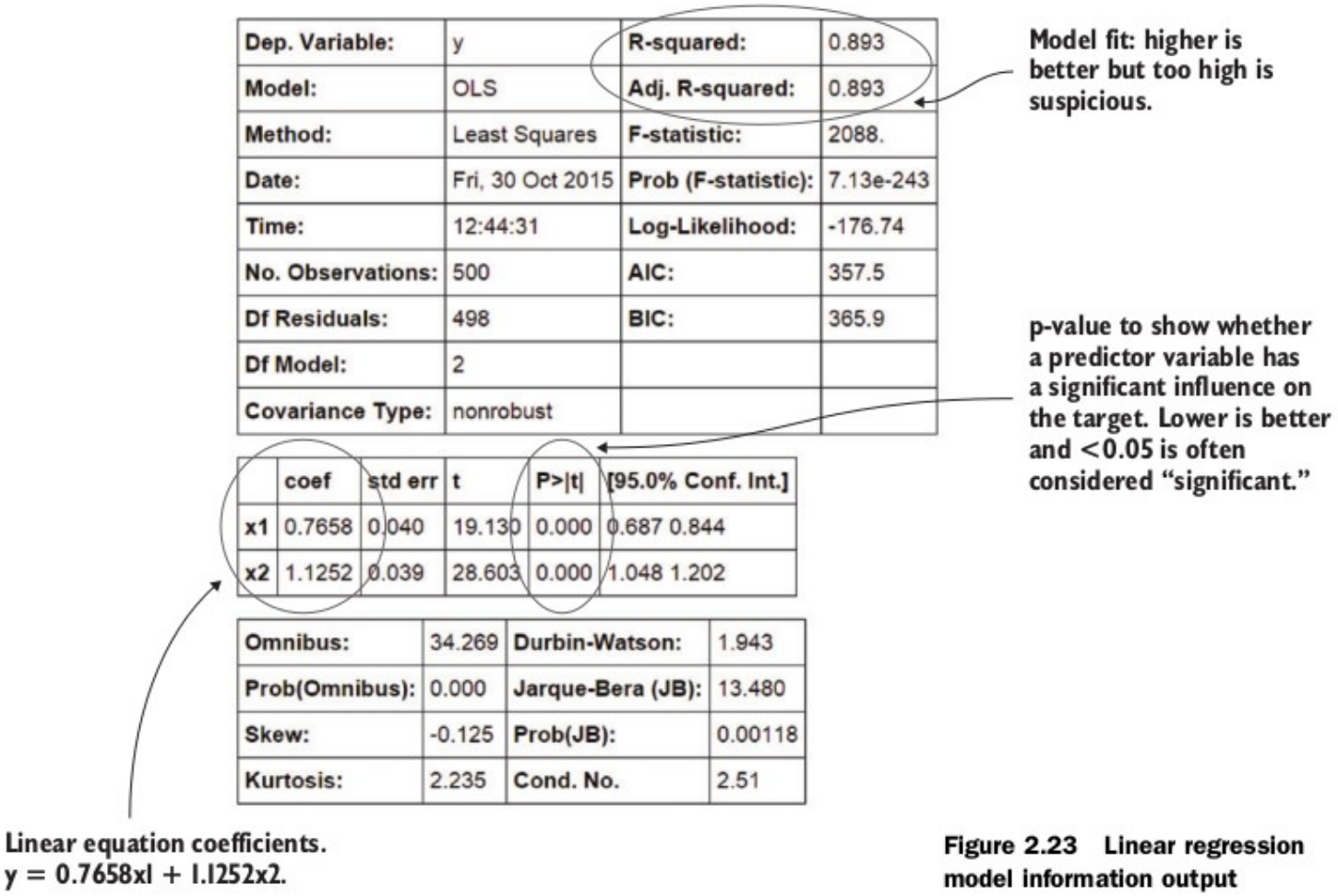
Imports required Python modules.

Creates random data for predictors (x-values) and semi-random data for the target (y-values) of the model. We use predictors as input to create the target so we infer a correlation here.

Fits linear regression on data.

Shows model fit statistics.

Evaluation



Dep. Variable:	y	R-squared:	0.893
Model:	OLS	Adj. R-squared:	0.893
Method:	Least Squares	F-statistic:	2088.
Date:	Fri, 30 Oct 2015	Prob (F-statistic):	7.13e-243
Time:	12:44:31	Log-Likelihood:	-176.74
No. Observations:	500	AIC:	357.5
Df Residuals:	498	BIC:	365.9
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
x1	0.7658	0.040	19.130	0.000	0.687 0.844
x2	1.1252	0.039	28.603	0.000	1.048 1.202

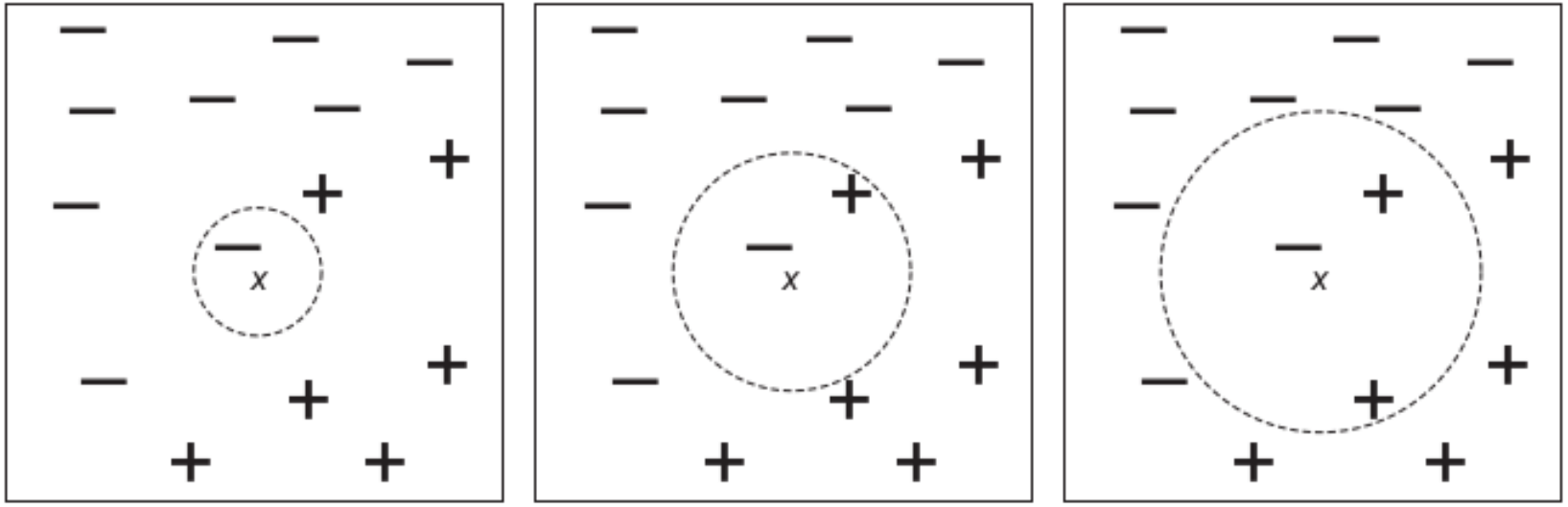
Omnibus:	34.269	Durbin-Watson:	1.943
Prob(Omnibus):	0.000	Jarque-Bera (JB):	13.480
Skew:	-0.125	Prob(JB):	0.00118
Kurtosis:	2.235	Cond. No.	2.51

Figure 2.23 Linear regression model information output

Evaluation

- Model fit—For this the R-squared or adjusted R-squared is used. This measure is an indication of the amount of variation in the data that gets captured by the model.
- Predictor variables have a coefficient—For a linear model this is easy to interpret. In our example if you add “1” to x1, it will change y by “0.7658”. It’s easy to see how finding a good predictor can be your route to a Nobel Prize even though your model as a whole is rubbish.
- Predictor significance—Coefficients are great, but sometimes not enough evidence exists to show that the influence is there. This is what the p-value is about.

Example: KNN Model



(a) 1-nearest neighbor (b) 2-nearest neighbor (c) 3-nearest neighbor

Code

Listing 2.2 Executing k-nearest neighbor classification on semi-random data

```
from sklearn import neighbors
predictors = np.random.random(1000).reshape(500,2)
target = np.around(predictors.dot(np.array([0.4, 0.6])) +
         np.random.random(500))
clf = neighbors.KNeighborsClassifier(n_neighbors=10)
knn = clf.fit(predictors,target)
knn.score(predictors, target)
```

Imports modules.

Creates random predictor data and semi-random target data based on predictor data.

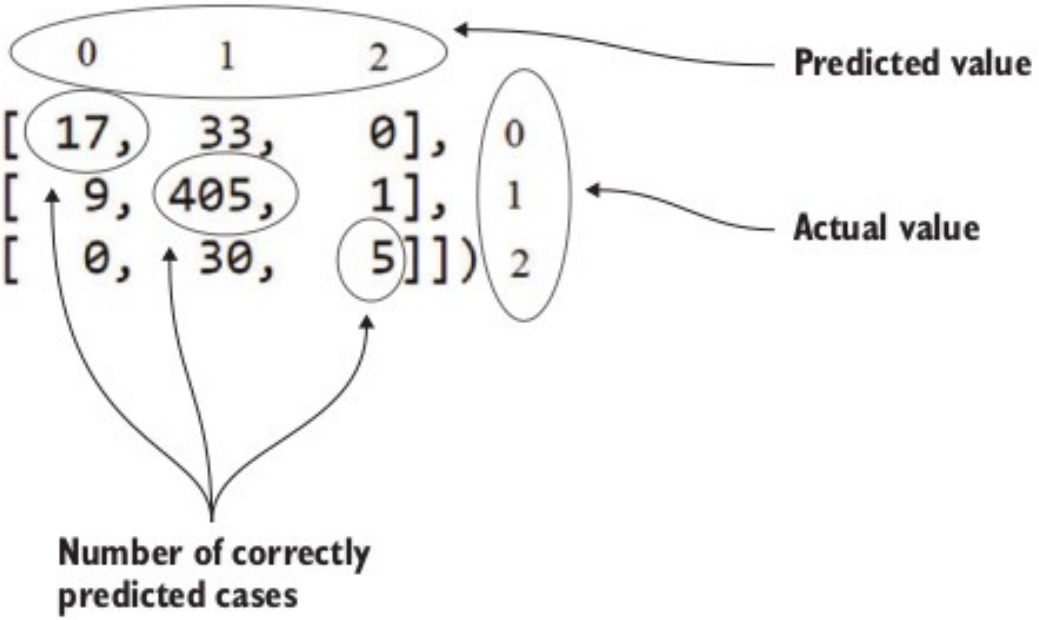
Fits 10-nearest neighbors model.

Gets model fit score: what percent of the classification was correct?

Evaluation

```
In [45]: metrics.confusion_matrix(target,prediction)

Out[45]: array([[ 17,  33,   0],
                [  9, 405,   1],
                [  0,  30,   5]])
```



Model diagnostic and comparison

- You'll be building multiple models from which you then choose the best one based on multiple criteria. Working with a holdout sample helps you pick the best-performing model. A holdout sample is a part of the data you leave out of the model building so it can be used to evaluate the model afterward. The principle here is simple: the model should work on unseen data. You use only a fraction of your data to estimate the
- model and the other part, the holdout sample, is kept out of the equation. The model
- is then unleashed on the unseen data and error measures are calculated to evaluate it.

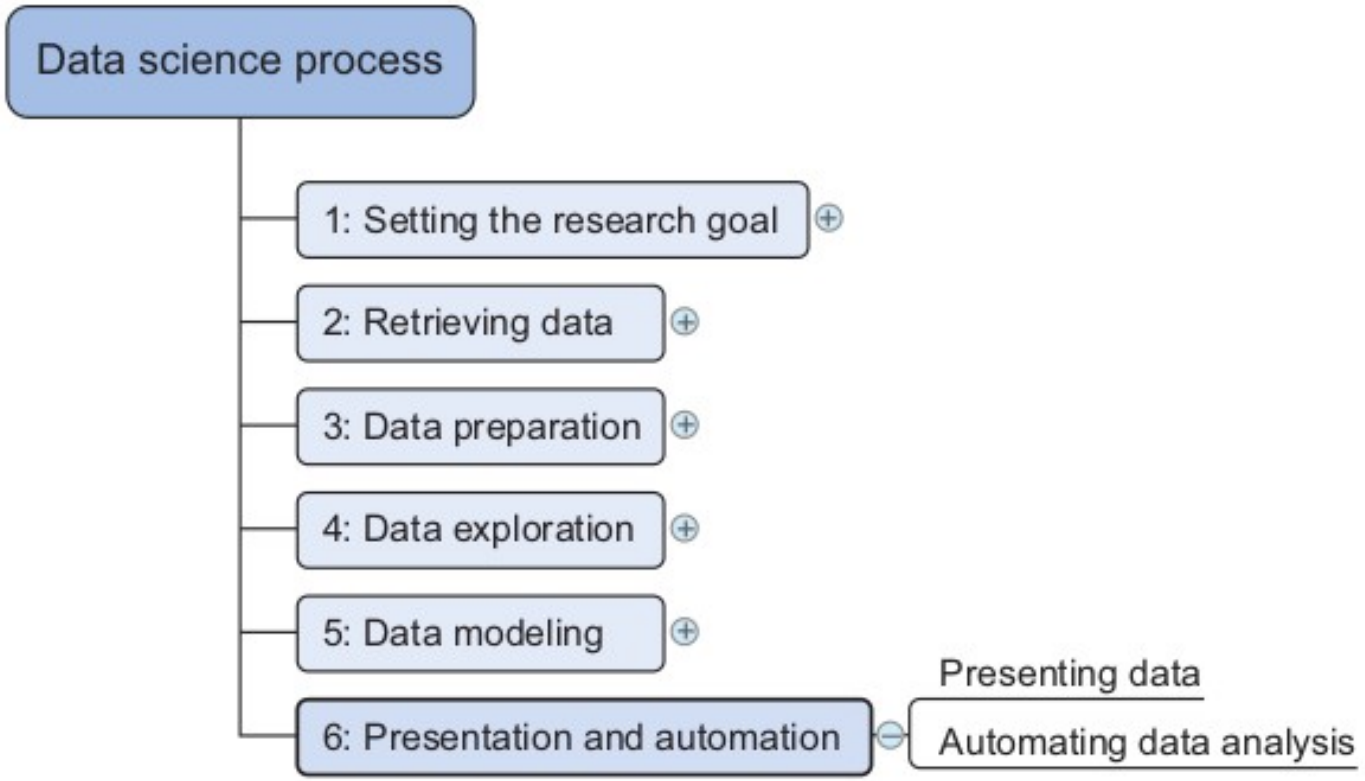
Cross Validation

	n	Size	Price	Predicted model 1	Predicted model 2	Error model 1	Error model 2
	1	10	3				
	2	15	5				
	3	18	6				
	4	14	5				
	…	…					
80% train	800	9	3				
	801	12	4	12	10	0	2
	802	13	4	12	10	1	3
	…						
	999	21	7	21	10	0	11
20% test	1000	10	4	12	10	–2	0
				Total		5861	110225

6. Presentation and automation

- After you've successfully analyzed the data and built a well-performing model, you're ready to present your findings to the world (figure).
- This is an exciting part; all your hours of hard work have paid off and you can explain what you found to the stakeholders.

6. Presentation and automation



Presentation

- Sometimes people get so excited about your work that you'll need to repeat it over and over again because they value the predictions of your models or the insights that you produced. For this reason, you need to automate your models.
- This doesn't always mean that you have to redo all of your analysis all the time.
- Sometimes it's sufficient that you implement only the model scoring; other times you might build an application that automatically updates reports, Excel spreadsheets, or PowerPoint presentations.
- The last stage of the data science process is where your soft skills will be most useful, and yes, they're extremely important.

Summary

- Setting the research goal—Defining the what, the why, and the how of your project in a project charter.
- Retrieving data—Finding and getting access to data needed in your project. This data is either found within the company or retrieved from a third party.
- Data preparation—Checking and remediating data errors, enriching the data with data from other data sources, and transforming it into a suitable format for your models.
- Data exploration—Diving deeper into your data using descriptive statistics and visual techniques.
- Data modeling—Using machine learning and statistical techniques to achieve your project goal.
- Presentation and automation—Presenting your results to the stakeholder and industrializing your analysis process for repetitive reuse and integration with other tools.

Thank you

@mitu_skillologies

/mITuSkillologies

@mitu_group

/company/mitu-skillologies

MITUSkillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

contact@mitu.co.in

tushar@tusharkute.com