

# Introduction to visualization NDBI046: Practical class 4



### **User Story**

- data warehouse
- \* We want to analyse the data to get an idea of the amount and origin of waste impact of waste management in the Czech Republic

### **Data Analyst roles:** \*

- Create *visualizations* of (various) data \*\*
- that it can be used for strategic decision making

Data engineers prepared and loaded the required waste management datasets to the

produced in each year and to be able to predict waste production in the coming years in order to target support and subsidy programmes that will address the (negative)

Interpret the results of the analysis and communicate them in clear and understandable terms so



### Tableau

- Data visualization software used for business inteligence purposes
  - Create *interactive and shareable visualizations* (i.e., graphs, charts, maps), dashboards, and stories
  - Supports various data sources including spreadsheets, databases, cloud services, and big data platforms
  - \* *Real-time data analysis*, i.e., supports visualization and analysis of data as it is being generated or updated
- Includes analytics capabilities such as predictive analytics, spatial analysis, and statistical functions, empowering users to make data-driven predictions and decisions
- User-friendly interface that enables users to drag and drop elements to create visualizations without requiring extensive programming knowledge
- \* Highly scalable, capable of handling large volumes of data
- Website: <u>https://www.tableau.com/</u>

### 



Show Me
on Name
AII)
lavní město Praha
hočeský kraj
homoravský kraj
arlovarský kraj
raj vysocina
borocký kraj
loravskoslezský k
lomoucký krai
ardubický kraj
lzeňský kraj
tředočeský kraj
stecký kraj
ínský kraj
AII)
014
015
016
017
ΠΤΩ
ure Names
unt of Waste Amo
Inning Sum of Was
< > > I



## **Prerequisite: Setting up Tableau Desktop**

- - Follow the installation
- Setting up PostgreSQL driver
  - \*
  - \* Copy the driver to the folder:
    - ~/Library/Tableau/Drivers (macOS)
    - \* /opt/tableau/tableau\_driver/jdbc (Linux)
    - C:\Program Files\Tableau\Drivers (Windows)

Download Tableau Desktop (14-day trial version) from <a href="https://www.tableau.com/">https://www.tableau.com/</a>

### Download the Java 8 JDBC driver from <a href="https://jdbc.postgresql.org/download/">https://jdbc.postgresql.org/download/</a>



### **Example 4.5: Connecting to a data source and selecting data**

- the webik.ms.mff.cuni.cz server
- - \* fact\_waste
  - \* dim\_years
  - \* dim\_regions
  - \* dim\_waste\_categories

Don't forget to save and name the project, e.g., ndbi046 \*

After installation, open Tableau Desktop, connect to the PostgreSQL data source on

\* Then select the datasets (i.e., tables) over which we will perform the visualization:



## **Example 4.5: Connecting to a data source and selecting data (Solution)**

- Open Tableau Desktop
  - \* When opened, a new project is automatically created (remember to save the project continuously)
- \* Select the data source *PostgreSQL* from *Connect* pane 1

  - Set and confirm the connection parameters (i.e., server, port, username, password) (2) \*
  - \* You are connected to the data source the available tables will be displayed in Table panel on the left 3

Installed Connectors (56)				Sort by Name (a-z) 🔻		
Alibaba AnalyticDB for MySQL Alibaba Data Lake Analytics Alibaba MaxCompute Amazon Athena Amazon Aurora for MySQL Amazon EMR Hadoop Hive Amazon Redshift Anaplan (deprecated) Apache Drill (deprecated) Azure Data Lake Storage Gen2 Azure SQL Database Azure Synapse Analytics Box	Google DriveHortonworks Hadoop HiveImpalaIntuit QuickBooks Online (deprecated)KyvosLinkedIn Sales Navigator (deprecated)MariaDBMarketing Cloud IntelligenceMarketo (deprecated)Microsoft SQL ServerMongoDB BI ConnectorMySQLODataOneDrive and SharePoint Online	Snowflake Spark SQL Teradata Vertica Web Data Connector (deprecated) Other Databases (JDBC) Other Databases (ODBC) Additional Connectors (34) ① Actian JDBC by Actian Actian ODBC by Actian Agiloft by Agiloft	Google Analytics 4 by Tableau Incorta by Incorta Jethro ODBC by Jethro Data Kyligence Connector by Kyligence MarkLogic by MarkLogic Ocient JDBC by Ocient OpenSearch by Amazon Oracle NetSuite by Tableau Palantir Foundry by Palantir Qubole Hive by Qubole Rockset by Rockset, Inc. Salesforce Marketing Cloud by Tableau SAP SuccessFactors by Tableau		PostgreSQL General Server webik.ms.m Port 5432 Database ndbi046 Authentication Username a	Initial S
Cloudera Hadoop Databricks Denodo (deprecated) Dremio Dropbox Esri Exasol Firebird 3 Google Analytics Google BigQuery Google BigQuery (JDBC)	Oracle Oracle Eloqua (deprecated) Pivotal Greenplum Databa PostgreSQL Presto Qubole Presto (deprecated) Salesforce Salesforce Data Cloud SAP HANA ServiceNow ITSM (deprecated) SharePoint Lists (deprecated)	Altinity Connector for ClickHouse by Altinity Inc Amazon DocumentDB by Amazon Amazon S3 by Tableau Bl Connector by Guidanz Inc Couchbase Analytics by Couchbase Analytics Data Cloud by Salesforce Data Virtuality JDBC by Data Virtuality Denodo JDBC by Denodo Technologies Exasol JDBC by Exasol	ServiceNow by Tableau Sharepoint Lists (JDBC) by Tableau SingleStoreDB JDBC by Singlestore Splunk by Tableau SQream DB by SQream Technologies Starburst Enterprise by Starburst Stratio Crossdata by Stratio BD Yellowbrick by Yellowbrick Data		Username koupil Password Optional I Require S	SL
	Alibaba Data Lake Analytics Alibaba MaxCompute Amazon Athena Amazon Aurora for MySQL Amazon EMR Hadoop Hive Amazon Redshift Anaplan (deprecated) Apache Drill (deprecated) Azure Data Lake Storage Gen2 Azure SQL Database Azure Synapse Analytics Box Cloudera Hadoop Databricks Denodo (deprecated) Dremio Dropbox Esri Exasol Firebird 3 Google Analytics Google BigQuery Google BigQuery (JDBC)	Alibaba Data Lake AnalyticsImpalaAlibaba MaxComputeIntuit QuickBooks Online (deprecated)Amazon AthenaKyvosAmazon Aurora for MySQLLinkedIn Sales Navigator (deprecated)Amazon EMR Hadoop HiveMariaDBAmazon RedshiftMarketing Cloud IntelligenceAnaplan (deprecated)Marketo (deprecated)Apache Drill (deprecated)Microsoft SQL ServerAzure Data Lake Storage Gen2MongoDB BI ConnectorAzure SQL DatabaseMySQLAzure Synapse AnalyticsODataBoxOneDrive and SharePoint OnlineCloudera HadoopOracleDatabricksOracle Eloqua (deprecated)DremioPostgreSQLDropboxPrestoEsriQubole Presto (deprecated)ExasolSalesforceFirebird 3Salesforce Data CloudGoogle AnalyticsSAP HANAGoogle BigQueryServiceNow ITSM (deprecated)Google Cloud SQLSingleStore	Alibaba Data Lake AnalyticsImpalaTeradataAlibaba MaxComputeIntuit QuickBooks Online (deprecated)VerticaAmazon AthenaKyvosWeb Data Connector (deprecated)Amazon Aurora for MySQLLinkedIn Sales Navigator (deprecated)Amazon EMR Hadoop HiveMariaDBOther Databases (JDBC)Amazon RedshiftMarketing Cloud IntelligenceOther Databases (ODBC)Anaplan (deprecated)Marketo (deprecated)Apache Drill (deprecated)Microsoft SQL ServerActian JDBC by ActianAzure SQL DatabaseMySQLActian JDBC by ActianAzure SQL DatabaseODataActian ODBC by ActianAzure Synapse AnalyticsODataActian ODBC by ActianBoxOneDrive and SharePoint OnlineAgiloft by AgiloftCloudera HadoopOracleAttinity Connector for ClickHouse by Attinity IncDatabricksOracle Eloqua (deprecated)Amazon S3 by TableauBremioPostgreSQLAttany DB by AmazonDrephoxPrestoData Cloud by SalesforceEsriQubole Presto (deprecated)Data Virtuality JDBC by Data VirtualityGoogle AnalyticsSAP HANADenodo JDBC by Denodo TechnologiesGoogle BigQueryServiceNow ITSM (deprecated)Exasol JDBC by Denodo TechnologiesGoogle Cloud SQLSingleStoreExasol JDBC by Exasol	Albaba Data Lake AnalyticsImpalaTeradataJethro ODBC by Jethro DataAlibaba MaxComputeIntuit QuickBooks Online (deprecated)VerticaKyligence Connector by KyligenceArnazon AthenaKyvosWeb Data Connector (deprecated)MarkLogic by MarkLogicArnazon Aurora for MySQLLinkedin Sales Navigator (deprecated)Ocient JDBC by OcientArnazon Atmacon EMR Hadoop HiveMarketing Cloud IntelligenceOther Databases (JDBC)Oracie NetSuite by TableauAnaplan (deprecated)Marketing Cloud IntelligenceOther Databases (ODBC)Oracie NetSuite by TableauAnaplan (deprecated)Marketing Cloud IntelligenceQuoble Hive by QuobleRockset by Rockset, Inc.Azure Data Lake Storage Gen2MongoDB BI ConnectorAdditional Connectors (34)ORockset by Rockset, Inc.Azure SQL DatabaseMySQLActian JDBC by ActianSalesforce Marketing Cloud by TableauSalesforce Marketing Cloud by TableauBoxOneDrive and SharePoint OnlineAttian (Donnector for ClickHouse by Attiant by Rockset, Inc.Salesforce Marketing Cloud by TableauDenodo (deprecated)PrestoOracleAttian OBC by ActianSalesforce Marketing Cloud by TableauDenodo (deprecated)PrestoQuoble Presto (deprecated)Marketo Single Store Data CloudDenodo (deprecated)PrestoCouchase Analytics by Couchbase AnajticsSplank by TableauBiomecore Di SubseroeSalesforceData Cloud by SalestoreStarburst Enterprise by StarburstDenodo (deprecated)SalesforceData Cloud by S	Alibaba Data Lake AnalyticsImpalaTeradataJethro DBB by Juthro DataAlibaba MaxComputeIntuit QuickBooks Online (deprecated)VerticaKylgence Connector by KylgenceAmazon AtternaKyvosWeb Data Connector (deprecated)MarkLagie ty MarkLagieAmazon Kertakoop HueMarketing Cloud IntelligenceOther Databases (DBC)OpenSearch by AmazonAmazon Kertakoop HueMarketing Cloud IntelligenceOther Databases (DBC)Oracle NetSuite by TableauAnazon Kertakoop HueMarketing Cloud IntelligenceOther Databases (DBC)Oracle NetSuite by TableauAnazon Kertakoop HueMarketing Cloud IntelligenceAdditional Connectors (34)OAzure Data Lake Stares Gen2MorgoDB BI ConnectorAdditional Connectors (34)OAzure SQL DatabaseMSQLActian JDBC by ActianSafesforee Marketing Cloud by TableauAzure SQL DatabaseOneDrive and SharePoint OnlineAgliot by AgliotSafesforee Marketing Cloud by TableauGoudera HadoopOracleOracle CongenetationSafesfore Marketing Cloud by TableauDataforksOracle Eloqua (deprecated)Amazon Document/DB by AmazonSafesforee Marketing Cloud by TableauDataforksOracle Eloqua (deprecated)Amazon Document/DB by AmazonSafesforee Marketing Cloud by TableauDataforksOracle Eloqua (deprecated)Amazon S3 by TableauSafesfore Marketing Cloud by TableauDataforksSafesforeSafesfore Marketing Cloud by SafesforeSafesfore Marketing Cloud by TableauDataforksSafesfore C	Albaba Data Lake Analytics     Impala     Teradata     Jethro OBBC by Jethro Data       Albaba MaxComputa     Intuit QuickBooks Online (deprecated)     Weth Ca     Kyligence Connector by Kyligence     General       Amazon Athena     Kyvos     Web Data Connector (deprecated)     MarkLogic by MarkLogic     General       Amazon Athena     Kyvos     Markelo (deprecated)     Other Databases (DBC)     OperSearch by Amazon     Weblik.ms.rn       Anazon Rodshift     Markelo (deprecated)     Other Databases (OBC)     Oracle NetSuite by Tableau     Port       Anazon Kotshift     Markelo (deprecated)     Markelo (deprecated)     Port     5432       Azare SQL Database     MySQL     Actian JDBC by Actian     Salesforce Markeling Cloud by Tableau     Port       Azare SQL Database     MySQL     Actian JDBC by Actian     Salesforce Markeling Cloud by Tableau     Salesforce Markeling Cloud by Tableau       Clouder Hadoop     Oracle Eloqua (deprecated)     Actian DBC by Actian     Salesforce Markeling Cloud by Tableau     Salesforce Markeling Cloud by Salesforce

\* If you do not have the *PostgreSQL Java 8 driver installed*, follow the instructions on page 4 (Prerequisite: Setting up Tableau Desktop)

		Tal	oleau - Book1 - Tableau license ex	pires in 10 days	
2		⊖- ndbi046			
sqL	Connections Add     vebik.ms.mff.cuni.cz   PostgreSQL     Database   ndbiO46     Table   III   dim_expenses_category   III   dim_regions   III   dim_waste_category   III   fact_expenses   III   fact_waste   III   student_table   III   New Custom SQL   III   New Table Extension			Drag tables here	
Sign In					
	O Data Source Sheet 1 ⊡ H H C				



### **Example 4.5: Connecting to a data source and selecting data (Solution)**

- Data selection
  - Drag and drop fact\_waste table from Table panel to the right area
  - Double click on fact\_waste table in the right area
  - A Drag and drop dim\_years table to the right area (Tableau automatically detects columns fact\_waste.year = dim\_years.year to perform an Inner Join)
     3
  - Drag and drop dim\_regions to the right area (dialog box appears)

<ul><li>Select</li></ul>	Inner Join <b>on the</b>	e fact_w	aste.reg	gion_ref =	dim_regions.	.region_id c	columns	5												
<ul><li>Drag and di</li><li>Perform</li></ul>	r <mark>op</mark> dim_waste_c <mark>n an</mark> Inner Join c	ategorie on the fa	es to the	e right are te.cate_re	a 6 f = dim_waste	e_categorie	s.cate_id	columns 7								Inne		Jeft	Pin Right	
	🖄 Tab	leau - ndbi046 - Table	au license expires in	10 days								<u>*</u>	🖄 Tableau - ndb	i046 - Table	eau license expires in	Dat Region R	a Source ef	=	5 Re	dim_ı gion Id
<ul> <li>★ → ➡ - □ - □</li> <li>Connections</li> <li>Add</li> <li>webik.ms.mff.cuni.cz</li> <li>PostgreSQL</li> <li>Database</li> <li>ndbi046</li> <li>Table</li> <li>p</li> <li>iii dim_expenses_category</li> <li>iii dim_regions</li> <li>iii dim_waste_category</li> <li>iii dim_years</li> <li>iii fact_expenses</li> <li>iii fact_waste</li> <li>iii student_table</li> </ul>	Correct fact_waste (ndbi046)	C	Need mo Drag tables here to rela	ore data? ate them. <u>Learn more</u>	Connection	Filters 0   Add		★ → ➡ → □       ↓         Connections       Add         webik.ms.mff.cuni.cz       PostgreSQL         Database       ndbi046         Table          Im dim_expenses_category       Im dim_regions         Im dim_waste_category       Im dim_years         Im fact_expenses       Im fact_expenses         Im student_table       Im student_table		<ul> <li>fact_was</li> <li>fact_wast</li> <li>fact_waste</li> </ul>	ste+ (no	de of 4 table	es. 0	dim_region dim_wast	ons te_category ts 6	Add new	join clause		7	
<ul><li>B New Custom SQL</li><li>P New Union</li></ul>	fact_waste    •   6 fields 622	rows	щ		100	→ rows ۞ ∨		😨 New Custom SQL	fac	et_waste		▼ 13 fie	elds 622 rows						100 → rc	)WS 🎲
$\beta_{db}^{h}$ New Table Extension	*     #     #     #       site     fact_waste     fact_waste     fact_waste       *     Region Ref     Year     Cate Ref	fact_waste	fact_waste Waste Amount	fact_waste Waste Per Capita				eas New Table Extension	Details	# fact_waste Region Ref	fac • Ye:	ŧ # t_waste fact_ ar Cate	waste fact_	waste	# fact_waste Waste Amount	# fact_waste Waste Per Capita	· dim_years Year (Dim Years)	# dim_regions <b>Region Id</b>	Abc dim_regions Region Name	d . <b>F</b>
	Image: Part of the second se	1         573,993.00           1         510,006.00           1         1,309,139.00           1         551,730.00           1         551,730.00           1         299,880.00	144,478.00 167,964.00 464,289.00 150,331.00 376,250.00 92,045.00	0.25171 0.32934 0.35465 0.27247 0.30074 0.30694					Table C		4 10 2 8 1 5	2014 2014 2014 2014 2014 2014	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	573,993.00 510,006.00 ,309,139.00 551,730.00 1,251,075.00 299,880.00	144,478.00 167,964.00 464,289.00 150,331.00 376,250.00 92,045.00	0.25171 0.32934 0.35465 0.27247 0.30074 0.30694	201 201 201 201 201 201 201 201	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	<ul> <li>4 Plzeňský kraj</li> <li>10 Kraj Vysočina</li> <li>2 Středočeský k</li> <li>8 Královéhradec</li> <li>1 Hlavní město F</li> <li>5 Karlovarský kr</li> </ul>	raj :ký kraj Praha
O Data Source Sheet 1 ⊡ <sub>+</sub> ⊞ <sub>+</sub> U	•							O Data Source Sheet 1 🔤 🕀							1				14 4 >	▶







### **Example 4.6: First Visualization – Bar chart**

- Create a visualization that allows us to identify in which \* year the Czech Republic had the highest waste production, regardless of its origin and type
  - Use a bar chart •





## **Example 4.6: First Visualization – Bar chart (Solution)**

- Bar chart \*
  - Create a new Worksheet
- - Drag and drop the dimension Year from Data panel to the chart as Columns shelf
  - Drag and drop Waste Amount from Data panel to the chart as Rows shelf 3
    - Tableau itself aggregates the data as a SUM

	lala		💮 Tableau - ndbi046 - Tableau license expires in 10 days	
A			s ↓ = ℓ • ℓ • T ↓ Standard • Standard • Standard •	宁 📑 Show Me
Data	anel	inns		
fact_waste+(nobic		I Rows		
Search	Filters	Sheet 1		
✓			Drop field here	
# Region Id	Marks			
<ul> <li></li></ul>	T Automatic 🔻			
Abc Cznace	Color Size Text			
✓				
	Detail			
# Cate Ref				
# Region Ref # Year				
Abc Measure Names		field	Drop field here	
<ul> <li>dim_regions</li> <li>Region Area</li> </ul>		here		
✓				
<ul><li># Population</li><li># Waste Amount</li></ul>				
# Waste Per Capita				
<pre># fact_waste (Count) # Measure Values</pre>				
Data Source Sheet 1	1			< < → → III

\* Tableau itself identifies the dimensions and measures in the data and displays them in the Data panel



## **Example 4.7: Column hierarchy and detailed interactive chart**

- \* distinction between municipal and industrial waste
  - ★ That is, create a column hierarchy Year → Region Name → Stapro
- The visualization should additionally:
  - Present population data for the year for each region
  - Appropriately visualize the Waste Per Capita measurement
  - Allow filtering of waste type based on the Cznace attribute

### Create a *visualization* that allows us to determine the *largest waste production*, by year, then by county, and finally by the





## Example 4.7: Column hierarchy and detailed interactive chart (Solution)

- \* Bar chart with column hierarchy
  - Create a new Worksheet
  - Drag and drop Waste Amount measurement from the Data panel to Rows shelf
  - Drag and drop Year, Region Name and Stapro dimensions to the Columns shelf
  - Select the Year dimension in the Data panel and create a hierarchy using the context menu Hierarchy → Create Hierarchy...
    - Name the hierarchy, e.g., YearRegionStaproHierarchy
  - \* Move (drag and drop) the Region Name and Stapro attributes into the newly created hierarchy YearRegionStaproHierarchy
    - \* Note that the hierarchy YearRegionStaproHierarchy can be expanded [+] and collapsed [-]

• •				⊕ т	ableau	- ndbi	046 -	Tablea	u lice	nse ex	pires	in 10	days																
$\bigstar  \leftarrow \rightarrow \frown \cdot \Box  \in$		·			<u>_</u> - (	) - []	0 ⊉	St	andard	•			Ì		2										Ļ		Show	Me	
Data Analytics <	Pages	iii Column	s 🚺	ear			Reg	ion Nar	ne		Sta	pro																	
<pre>   fact_waste+ (ndbi046) </pre>		≣ Rows		UM(Wa	ste Amo	ount)																							
Search ♀ ♥ IIII ▼ Tables	Filters	Sheet	2				3																						
dim_regions								21	24.4			Yea	r / Re	egion	Name	/ Sta	pro							24	015				
<ul><li># Region Id</li><li>Abc Region Name</li></ul>	Marks		Hlavní Jihoč	Jihon	Karlov	Kraj	Králov	Libere	Mora	VOlom	o Paro	du Plze	ns Sti	ředoÚ	Ísteck	Zlínsk	Hlavi	ní Jiho	će Jiho	om Kar	lov K	raj K	rálov	Libere	Mora	av Olon	no Par	du Plz	
dim_waste_category	🛯 Automatic 💌		měst ský k	. oravs	arský	Vyso	éhra	cký k.	skosl.	. ucký	bick	ý ký k	raj čes	ský y	ý kraj	ý kraj	měst	: ský l	orav	/sars	ký Vy	'so é	hra	cký k.	. skos	I ucký	i bick	(ý., ký	
# Cate Id		15M																											
Abc Cznace Abc Stapro	Color Size Label																												
i dim_years																													
# Year (Dim Years)	Detail Tooltip	ゼ 10M																											
fact_waste		nom																											
# Cate Ref		ite A																											
# Region Ref		Was																											
# Year		5M																	_										
dim regions																													
# Region Area																										Ι.			
fact_waste		014																											
# Population											; ;																		
# Waste Amount				unu	un n	nună	nună	unn;	nuna	Punu		unu unu	nun	0 OL	no or	unn;	unu			nuna	un o	00		un o	unu		un o	unu unu	
# Waste Per Capita			e kor ovéh e kor	e kor	e kor ovél	e kor ovéł	e kor ovéł	e kor ovél	e kor	e kor	kor	ovél e kor	ové! kor	ovél	e kor ovéł	e kor ovél	kor	e kor	e kor	ovél e kor	ové! e kor	ovél	e kor ovéł	e kor ovél	kor	e kor	ové! e kor	ovél e kor	
# fact_waste (Count)			dnik ukce	ukce	ukce dnik	ukce dnik	ukce dnik	ukce dnik	ukce	ukce dnib	ukce	dnik ukce	dnik ukce	dnik	dnik	ukce dnik	ukce	ukce :	ukce	dnik ukce	dnik ukce	dnik	dnik	ukce dnik	ukc(	ukce .	dnik ukce	dnik ukce	
# Measure values			e po orod	orod	orod	orod e po	orod e po	or od e po	or od	orod	orod	e po orod	e po	e po	e po	or od e po	prod	e po	e po	e po orod	e po orod	e po	e po	orod	orod	e po	e po orod	e po prod	
			ová j dukc ová j	ová l	ová I dukc	ová I dukc	ová   dukc	ová I dukc	ová   huko	ová l	ová	dukc ová I	dukc ová I	, dukc	ová   dukc	ová I dukc	ová	ová I	ová	dukc ová j	dukc ová j	Jukc	ova   dukc	ová I dukc	ová	ová I	dukc ová j	dukc	
			Celk	Celk	Celk	Celki	Celki	Celk	Celk	Celk	Celk	Pro( Celk(	Proc	Proc	Celk	Celk	Celk	Celk(	Celk	Celk	Proc	Proc	Proc	Celk	Celk	Celk	Celk	Celk	
							-					10			-					1.0	10		-	-					
Data Source Sheet 1 Sh	heet 2																												
0 marks 1 row by 140 columns SU	UM(Waste) 26,967		_																						M	< > >			



## **Example 4.7: Column hierarchy and detailed interactive chart (Solution)**

- Adding additional information to the chart using Marks card \*
  - Represent the Population measurement as a label
  - Trag and drop the Population measurement from the Data panel to the Label in Marks card 1
     2 Open the context menu of the SUM(Population) in Marks card and select Measure (Sum) → Average 3 \* To visualize Waste Per Capita, apply a background color to each column
    - Trag and drop the Waste Per Capita measurement from the Data panel to Color in Marks card 4 5





## **Example 4.7: Column hierarchy and detailed interactive chart (Solution)**

- Addition of filtering the underlying data for visualization \*
  - Trag and drop the Cznace attribute from the Data panel to Filters shelf
    - Dialog box appears
    - \*
  - \* Open the context menu of the Cznace filter and select Show Filter to conveniently adjust the filtering using the 3 checkbox



Selected values of Cznace attribute can be specified, e.g., 'Stavebnictví', and confirm the selection 2

						14 14	Tablea	u - ndb	oi046 - T	ableau	license	expire	es in 10	days						
				<u></u>			<u>/</u> •	0 -	T∤	Fit W	/idth	•	8,11 ▼		ಹ					cne
	Data Analytics <	Pages	ii	i Columns	5	E Ye	ear		E R	egion Nai	me									
	🍃 fact_waste+ (ndbi046)			Rows		SUM	(Waste Ar	nount)												
on Top			3																	
Lise all		Filters		sheet	2															
	lables	Cznace										Yea	r / Reaid	on Name						
nažďování, sb	dim_regions					20	14		2	2015		reu	2016	ò		2	2017			2018
	# Region lu	Marks																		
	<ul> <li># Cate Id</li> </ul>	Automatia																101	not	
	Abc Cznace	Automatic		51.	12													100	то <u>о</u> т'	
ého vz	✓	<b>:</b> 6	Т	510	51,0			207										~		
	# Year (Dim Years)	Color Size	Label		1,25			62,5	~			32								
	fact_waste			4N	1-			1,2	,563			72,7			4					
	# Cate Ref	Detail Tooltip		nt					,173			1,27			,55,				729	
Exclude	# Region Ref	SUM(Wast	te Pe	nom					с-Î	60		972			,286	477	<u>م</u>		84,7	9
Exclude	✓ ♣ YearRegionStaproHierarc	T AVG(Popu	latio	A 3V	0		9,727			5,2(		176,	437			180,	7,41		1,1	4,37
	# Year			Was	0,67		,219			1,21		1,1	211	~~~		Ļ,	L,20			1,20
	Abc Region Name			20	1,17		۵ ۲	3			721		H .	638				37		
	Abc Stapro			210			19				20,7			577				50,9		
							130		2		1,3							80		
	# Region Area			1N	1	,006			9,5(				187	- 8		664			,019	
	✓ ■ fact waste					510			50				509,	- 11		508,			509	
	+ Population			ON	1															
ОК	# Waste Amount		-		 	a a	ם ס ס	<u> </u>	<u>.</u> . <u>.</u>	<u>a</u> . <u>a</u> .	a. a.	ים	ם ס	<u>.</u> 	aj aj	<u>a</u> . <u>a</u> .	<u>a</u> . <u>a</u> .	Taj.	ים סיים	<u>.</u> 
	# Waste Per Capita				ký ki	X X X	x x x x x x	K X	x x x	Ϋ́Ϋ́Υ	ký ki ký ki	ký ki ký	x x x x	ν γ <sup>×</sup> × <sup>×</sup>	ký ki	k k V	ký ki ký ki	ký ki	ký ki ký ki v ki	ký ki
	<pre># fact_waste (Count)</pre>				očes	/ars idec	lezs ubic ŏčes	líns	/ars	lezs ubic	očes	)čes /ars	idec	ubic	líns očes	/ars idec	lezs ubic	očes	Jčes Jars	lezs
	# Measure Values				Jiho	arlov éhra	skos ardı ředo		arlov éhra	skos	ředo	Jiho	éhra skos	ardı ředo	Z	arlov éhra	skos ardı	ředo	Jiho arlov	skos
							St		álov K	Frave	St	X	álov	St P		alov K	F	St	X Volie	ravs
						Y	Mo		X	Mo			A No			X	Mo		Х	Mo
	O Data Source Sheet 1 Sh	eet 2 🖳 🕂	Π.																	
	70 marks 1 row by 70 columns SUM(	Waste Amount): 80.41	3.219																	



### **Example 4.8: Underlying data**

- \*
  - Make a selection of relevant data •
  - Display the underlying data \*

Obtain the *underlying data* from the visualization in Example 4.8, but only for 2018

14

## **Example 4.8: Underlying data (Solution)**

- Selection of relevant data \*

  - \* Duplicate the Worksheet with the chart from Example 4.7 using Worksheet context menu Display visualization by year (i.e., collapse Year column)
  - Select the bar for '2018' and make the year selection using Keep Only from the context menu \* Or select simultaneously the years '2014', '2015', '2016', and '2017' and perform Exclude





## **Example 4.8: Underlying data (Solution)**

- Displaying the underlying data \*
  - Select the bar for '2018' and choose View Data... from the context menu 1
    - \* The *View Data* appears

  - To display all non-aggregated data, select Full Data from the Tabs panel 4

	View Data: Sheet 2 (2	?) (1 mark)					×		
	Tabs	Summary 1 row 4 field	ds			+ Show Fields	🖵 Download		
Comparison of the second	Summary 3	# fact_waste	# fact_waste	View Data: Sheet 2 (2)	(1 mark)				
$\widehat{\square}  \overleftarrow{\leftarrow} \rightarrow \stackrel{\frown}{\frown}  \overleftarrow{\square}  \overleftarrow{\leftarrow} \rightarrow \stackrel{\frown}{\frown}  \overleftarrow{\square}  \overleftarrow{\leftarrow} \rightarrow \stackrel{\frown}{\frown}  \underbrace{\square}  \overleftarrow{\frown} \rightarrow \stackrel{\frown}{\frown}  \overleftarrow{\frown} \rightarrow \stackrel{\frown}{\frown}  \underbrace{\square}  \overleftarrow{\frown} \rightarrow \stackrel{\frown}{\frown}  \overleftarrow \rightarrow \rightarrow \stackrel{\frown}{\frown}  \rightarrow \rightarrow \rightarrow \stackrel{\frown}{\frown} \rightarrow \stackrel{\frown}{\frown}  \rightarrow \rightarrow \rightarrow \stackrel{\frown}{\frown}  \rightarrow \rightarrow \rightarrow \stackrel{\frown}{\frown}  \rightarrow \rightarrow$		Year	Avg. Popula	Tabs <	<b>Full Data</b> 54 rows 5	fields		[+ \$	Show Fields 🛛 🖓 🖓
Search $\rho$ $\gamma$ $\blacksquare$ Filters Cznace		2018	7	Summary	#	Abc	#	#	#
Tables     Cznace     (All)       Image: marked bit in the second bit in the seco				🔲 Full Data 🛛 🛓	fact_waste	fact_waste	fact_waste	fact_waste	fact_waste
# Region Id     Doprava a skladování       ∨ ⊞ dim_waste_category     18M					Year	Cznace	Population	Waste Amount	Waste Per Ca
# Cate Id     Marks       Abc     Cznace       16M     16M       16M     16M					2018	Stavebnictví	582,601.00	364,688.00	
<ul> <li>✓ III dim_years</li> <li>III Zemědělství, lesnic</li> <li>✓ III Zemědělství, lesnic</li> <li>✓ Zpracovatelský prů</li> </ul>					2018	Stavebnictví	509,019.00	194,187.00	
# Cate Ref     000     000     000     000       # Region Ref     Detail     Tooltip     12M     12M					2018	Stavebnictví	1,360,998.00	553,699.00	
YearRegionStaproHierarc # Year I AVG(Populatio) AVG(Populatio) AVG(Populatio) 4 10.55 10.54 10.55 10					2018	Stavebnictví	550,688.00	240,109.00	
Abc     Region Name       Abc     Stapro   Abc Stapro Abc Stapro Abc Access and a stapro Abc Abc Access and a stapro Acc					2018	Stavebnictví	1,301,135.00	4,210,257.00	
Accomeasure warres     > 8M     Year:     2018       > Im dim_regions     Avg. Population:     767,519       # Region Area     Waste Amount:     17,279,304					2018	Stavebnictví	295,285,00	85,705,00	
fact_waste     6M     Waste Per Capita: 19.54       # Population     6M     4					2018	Stavobnictví	1 204 346 00	685 523 00	
# Waste Amount     4M       # Waste Per Capita     4M	→ rows <b>čõ</b>	3			2018	Stavebhictvi	1,204,340.00	003,323.00	
# fact_waste (Count)     2M       # Measure Values     2M					2018	Stavebnictví	519,125.00	456,369.00	
OM					2018	Stavebnictví	582,860.00	384,808.00	
					2018	Stavebnictví	640,909.00	570,279.00	
Data Source Sheet 1 Sheet 2 (2) Ut Ht Ut				10,000 → rows 🔅	2019	Stavobnictví	1 194 720 00	2 209 572 00	

\* The View Data Sheet contains one row of aggregated data corresponding to the selected bar (i.e., Summary) 3





### **Example 4.9: Dual axis chart**

- Create a single visualization that allows for analysis to compare Waste Per Capita and Waste Per Area over regions with respect to Region Area and Population
  - Calculate Waste Per Area as a ratio of Waste Amount and • Region Area
  - Create a dual axis chart \*
    - Plot the range of Waste Per Capita measurement values on the left Y-axis
    - Plot the range of Waste Per Area calculated field values on the right Y-axis
    - Visualize the values of Waste Per Capita and Waste Per Area with *different colors and shapes* 
      - \* Reflect the *magnitude of the value by the size* of the displayed shape



Waste Per Capita 7.00 20.00 30.00 40.00 48.41 Naste Per Area 629 20.000 60,000 80,000 100,000 124,461 Waste Per Capit

17

### **Example 4.9: Dual axis chart (Solution)**

- \* Adding calculated field Waste Per Area
  - Select the Waste Amount measurement in the Data panel
    - \* In the context menu, select Create  $\rightarrow$  Calculated Field... 2
  - \* A dialog box appears that allows us to name and define a new calculated field
    - Name the new field Waste Per Area
    - The field is calculated using the formula [Waste Amount] / [Region Area] (note that attribute names are entered in square brackets)
    - \* After hitting Apply, a new field appears in the Data panel (if the formula is valid)
    - Confirm the creation by hitting OK

••	•			🕸 Tablea	u - ndbi046 - 1	Tableau license exp	ires in 10 days	;		
A	$\leftarrow \rightarrow \multimap \cdot \square \square$				0 - T Z	Standard -		ಹ್	宁 📑 Show Me	
Data	Analytics <	Pages	iii Columns							
🖯 fa	ict_waste+ (ndbi046)		≣ Rows							
Seare										
Table	· · · · · · · · · · · · · · · · · · ·	Filters	Sheet 4							
→ Ⅲ ( # → Ⅲ (	dim_regions Region Id Iim waste category	Marks					Dr	op field here		Add to She
# Abc	Cate Id Cznace Jim_years	T Automatic								<b>Rename</b> Hide
# • ■ f #	Year (Dim Years) fact_waste Cate Ref Bogion Bof	Color Size Text								Create Convert to
# /	YearRegionStaproHierarc Year Region Name		Drop							Convert to Change D Default Pi
Abo	Stapro		field				Dr	op field here		Geograph
× ⊞ ( # ~ ⊞ f	lim_regions Region Area		Tiere							Group by Folders
# # # # 1	Waste Amount Waste Per Capita									Replace R Describe.
# /	Measure Values									
0 Data	Source Sheet 1 She	eet 2 Sheet 2 (2) Sheet 4								









### Example 4.9: Dual axis chart (Solution)

- \* Dual axis chart
  - \* Create a new Worksheet
  - Drag and drop Region Name attribute from the Data panel to Columns shelf
  - Drag and drop Waste Per Capita and Waste Per Area measurements from the Data panel to Rows shelf (see visualization is on the left)
     3
  - Open the context menu of the Y-axis in the bottom chart and choose Dual Axis
- \* Adjust the data points to make their size proportional to the value of the corresponding measures
  - \* Open Marks shelf for SUM(Waste Per Capita) and drag and drop the Waste Per Capita measurement from the Data panel to Size in Marks shelf
  - Open Marks shelf for SUM(Waste Per Area) and drag and drop the Waste Per Area measurement from the Data panel to Size in Marks shelf
- Open the drop-down menu in SUM(Waste Per Area) Marks shelf and change the shape from 'Automatic' to 'Square' 8







Add Reference Line

# Region Area

fact\_waste

# Populatio

Waste Amoun

# Waste Per Capita

Waste Per Area

# fact\_waste (Count)
# Measure Values

🖯 Data Source Sheet 1 Sheet 2 Sheet 2 (2) Sheet 4 🖳 🕀 🗸

28 marks 1 row by 14 columns SUM(Waste Per Area): 151,904



### **Example 4.10: Scatter plot**

- Create a visualization that can be used to determine the relationship between measurements (average) Population and (average) Waste Amount
  - Use a more detailed breakdown of the points, specifically break them down using **Region Name**

 Create a linear regression curve of waste on population

**Region Name** 1400K Hlavní město Praha Jihočeský kraj Jihomoravský kraj Karlovarský kraj 1200K Kraj Vysočina Královéhradecký kraj Liberecký kraj 1000K Moravskoslezský kraj Olomoucký kraj Avg. Waste Amount Pardubický kraj 800K Plzeňský kraj Středočeský kraj Ústecký kraj Zlínský kraj 600K 400K 200K ОК 1400K ОK 800K 1000K 1200K Avg. Population



### **Example 4.10: Scatter plot (Solution)**

- Scatter plot
  - \* Create a new Worksheet
  - Drag and drop Population measurement from the Data panel to Columns shelf

  - Drag and drop Waste Amount measurement from the Data panel to Rows shelf
  - \*
  - Trag and drop Region Name onto Color item in Marks shelf (to divide the points in the chart based on the region and color them)
     6
  - Open Shape in the Marks shelf and select

		🕸 Tableau	Tableau license expires in 10 day	ys	
		li 🧗 🖉 Z	Standard 💌 🗄 🛄 🗸 👾	с <mark>о</mark>	宁 📑 Show Me
Data Analytics < Pages	iii Columns	AVG(Population)			
🗟 fact_waste+ (ndbi046)	<b>≔</b> Rows	AVG(Waste Amount)			
Search $\rho$ $\nabla$ IIII - Filters					
Tables	Sheet 5	4			
✓				0	Filter
# Region Id     ✓      # dim_waste_category     # Cate Id	350К				Show Filte Apply to V
Abc Cznace	) T 300K				Format ✓ Include in
# Year (Dim Years)       ✓ Im fact_waste       # Cate Ref	ail Tooltip 250K				Dimension Attribute
# Region Ref     ✓	aste Amou				Discrete ✓ Continuor
Abc Region Name Abc Stapro Abc Measure Names	М. бок 150К				Edit in Sh
<ul><li>✓ I dim_regions</li><li># Region Area</li></ul>					Add Table Quick Tab
<ul> <li>main fact_waste</li> <li># Population</li> </ul>	100K				Remove
# Waste Amount     # Waste Per Capita     # Waste Per Area	50K				
<ul> <li># fact_waste (Count)</li> <li># Measure Values</li> </ul>	ок	100K 200K 300K	400K 500K 600K 700	к 800К	
		A	vg. Population		
O Data Source         Sheet 1         Sheet 2         Sheet 2           1 mark         1 row by 1 column         SUM of AVG(Population): 757	(2) Sheet 4 Sheet 5 ,850				





### **Example 4.10: Scatter plot (Solution)**

- Adding a linear regression curve •
  - Open the Analytics panel
  - Drag and drop Trend Line to the chart area (as 'linear') \*







### **Example 4.11: Pareto chart**

- Create a visualization that captures the frequency of amounts of waste production \*
- Then, adjust the visualization to identify which waste components (i.e., small or large production) contribute the most to the total amount of waste produced
  - Construct a *running total* of that measure and visualize it as a *line graph* •



### Solution with the second se





## **Example 4.11: Pareto chart (Solution)**

- Creating a histogram

  - Waste Amount 4







### ... more in Data visualization (NDBI042)

- Lecturer: David Hoksza
- Syllabus
  - Introduction to data representation and visualization •
  - Visualization types and principles
  - Data visualization on the Web \*
  - Multi-dimensional data visualization PCA \*
  - Multi-dimensional scaling •
  - Graph data visualization \*





### Web: <u>https://bioinformatics.cuni.cz/cusbg/members/hoksza/courses/visualization.html</u>



### References

### **Tableau**

- Download: •
  - https://www.tableau.com/ \*
- **Documentation:** \*
  - \*
- \* Tutorials:
  - \* tutorial-home.htm

### https://help.tableau.com/current/pro/desktop/en-us/gettingstarted\_overview.htm

### https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-

