



**UNIVERSITA' DEL SALENTO**

**DEPARTMENT OF ENGINEERING FOR INNOVATION  
MANAGEMENT ENGINEERING**

---

*Report Data Contest  
Database Development for PNRR Application  
Data Management Course*

**STUDENTS:**  
*Foggetti Angelo  
Nesca Gianvito  
Zarrella Davide*

---

Academic Year 2022- 2023



## SUMMARY

INTRODUCTION .....	4
Scope .....	4
Focus on the context: PNRR and databases .....	4
Method & Tools .....	6
DATASET DESCRIPTION .....	8
THE ENHANCED ENTITY-RELATIONSHIP MODEL.....	24
THE RELATIONAL MODEL .....	28
DATABASE DEVELOPMENT .....	31
Requested queries .....	34
Additional queries .....	41
REFERENCES .....	47

# INTRODUCTION

## Scope

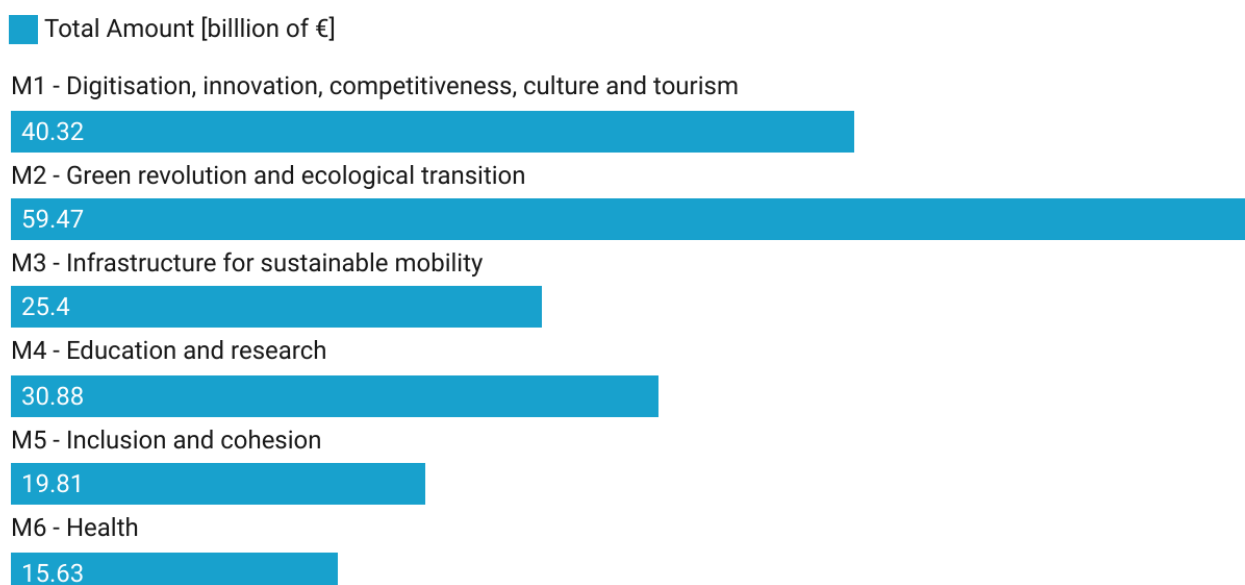
This report aims to analyze different sets of data, inherent to the planning, enactment, and localization of projects about PNRR, to transform them appropriately into a database to support the decision-making process.

## Focus on the context: PNRR and databases

As said above, this report aims to build a database capable of containing PNRR data, that can be used to obtain valuable information for the decision-making process. However, before going into the core of the report, it is preferable to frame the context and understand what the PNRR is, what a database is and how these two concepts can intersect with each other to achieve the goal.

The PNRR, acronym of “Piano Nazionale di Ripresa e Resilienza”, is the strategic document that the Italian Government has prepared to obtain the funds of the Next generation EU Program (NGEU), the 750-billion-euro package agreed by the European Union. The Plan was presented on 30 April 2021 to the European Commission and was approved on 13 July 2021.

It foresees investments amounting to 191.5 billion euros over a period of 5 years (until 2026) and, as the Figure 1 shows, it is divided into 6 Missions, each of which is further divided into Components, Measures and Sub-measures.



**Figure 1.** PNRR Missions  
Revisited by PNRR, 2021

Its main objectives are:

- relaunching the country after the pandemic crisis, stimulating an ecological and digital transition;
- encouraging structural change in the economy, starting with the fight against gender, territorial and generational inequalities.

However, the PNRR is not only a traditional investment program, but it is a transformative project, where the investments of resources are accompanied by a substantial package of reforms necessary to overcome the historical barriers that have hindered the development of public and private investment, creating unsatisfactory employment levels, especially for young people and women (PNRR, 2021; [www.governo.it](http://www.governo.it)).

In the case of the PNRR, being made up of a set of projects, it generates a large amount of data which must be organized in the best possible way, to be consulted and queried. This data is stored in "containers" called Databases, to which all programs and users can access simultaneously and with simple languages. Obviously, these considerations can be extended to any area because we live in an information-based society, and these have grown very rapidly in recent years so much so that everything we do with our devices can be seen as a set of data (photos, videos, emails, calls, etc.).

We define a database as a collection of structured data, designed to be used in different application and by different users. For their management and verification of their consistency, special software systems are required, called DBMS (Data Base Management System) which deal with the storage, organization, and maintenance of data.

In the context of database design, four independent and consecutive levels are distinguished.

- Requirements specification and analysis.
- Conceptual Design
- Logical Design (Data Model Mapping).
- Physical Design.

With the first level we can collect all the information we need to develop the database: all these requirements are documented in detail and transformed into a conceptual design, that is a graphic conceptual form at a high level of abstraction. One of the most used conceptual data models is the Entity-Relationship Model (or Enhanced Entity-Relationship Model), a popular high-level diagram in which different entities are connected to each other through relationships, with their own cardinality (1:1, 1:N, M:N) and a total or partial participation constraint (Elmasri and Navathe, 2015). The second step foresees the development of the logical model: among all these, the Relational model is probably the most widespread in databases. It is so named because each table of the databases is a relation, understood in the sense of mathematical logic. Values taken from a row of the table are called tuples or records. In the relational model the constraints within the relations or two different relations assume fundamental importance, so much so that we can distinguish the:

- Primary key, that is a subset of attributes whereby each value uniquely identifies each tuple of the Relation. It is important to note that the values of a primary key must always be entered (it is also known that it does not admit NULL values) (Entity integrity constraint); and the:
- Foreign Key, that is an attribute (or set of attributes) that links one relation to another, imposing that there is a correspondence between the attributes common to the two tables. In particular, the referential integrity constraint between R1 and R2 dictates that if F is the foreign key of R2, then all F values that appear in R2 must appear as primary key values in R1 (Referential integrity constraint).

Whereas columns mean attributes and rows mean tuples or records, you can list 7 different conversion rules that allow you to move from the conceptual model to the logical one in a relational database.

1. Each entity becomes a table.
2. Entity attributes become table columns.
3. Columns inherit attributes characteristics.
4. The key attribute of the entity becomes the primary key of the table.
5. If the association is 1 to N, on the N side a column is added, corresponding to the primary key of side 1. This column is the foreign key of the relation.
6. If the association is 1 to 1, you can choose where to add the column.
7. If the association is M to N, a third table is added, which contains the keys of the other two tables ([www.villaggioinformatico.it](http://www.villaggioinformatico.it)).

The final stage is physical design, during which further specifications are provided for storing and accessing the database. The database design is implemented, populated with actual data, and continuously maintained to reflect the state of the *miniworld* (Elmasri and Navathe, 2015).

## Method & Tools

Developing a functioning database capable of collecting and querying PNRR data required the performance of several sequential steps.

After downloading the files from the site, the first step we took was the analysis of the data contained in the various tables; therefore, after loading all 15 files in the Microsoft Excel worksheets, we tried to understand the meaning of the attributes, also helping us with the metadata, which can be downloaded from the same link that leads back to the tables. This process, better known as *Data Exploration*, has allowed us to discover the characteristics of the data, their nature, and the positive and negative peculiarities of each dataset. The first chapter, in fact, called *Data Description*, contains the description of each dataset, and shows its characteristics in terms of consistency, accuracy, quality and completeness (for example, as regards this last aspect, for each column of each dataset it is been counted, using the Excel filter tool, the number of null values).

After studying each dataset, understanding the topic covered, the second step was the development of the *EER Model*: to do this we expanded the perspective and analysed the datasets no longer as standalone items but as a more complex and interactive system, understanding through the common attributes found in each table the relationships that could exist between the different datasets. Doing this, we understood which attributes and which datasets could be an entity, depending on whether for each of them it was possible to trace a key attribute that distinguished each entity, and we identified the relationships between the different entities, their cardinality relationship, and the constraints of participation, total and partial.

The last act before developing the database was the transition from the conceptual model to the logical model. Indeed, starting from the EER, we developed the relation database schema, identifying all the Relation Schemas and attributes, mapping all the relationships between the entities of the EER and identifying the primary and the foreign keys. To draw both the conceptual and the Relational Model, we used the open source draw.io graphic design software.

As for the physical development of the database, it was built through MySQL Workbench and its creation required two different steps, which are the creation of the database structure and the loading of data to populate it. The first step returned the schema database, that is the description of the tables, specifying the attribute data types, the domains and the relations between all the tables due to the primary and foreign keys. More complicated was the second step: to clean the data, in fact, we used several tools, including Excel, Power BI and MySQL Workbench, which allowed us to delete unnecessary columns, delete duplicates, intersect or merge two or more datasets, rename the null values and thus get clean tables to load into the database to populate it, obtaining the extension (database state) of the previously created schema. To populate the databases we used the website [products.aspose.app](https://products.aspose.app) that allows you to create, starting from excel files, the script of the command INSERT INTO TABLE. Once we created the database on MySQL Workbench, we used this same tool to query the database through queries to analyze the data from a different and wider perspective. The Figure 2 summarizes schematically all the steps we have performed and the tools we have used.

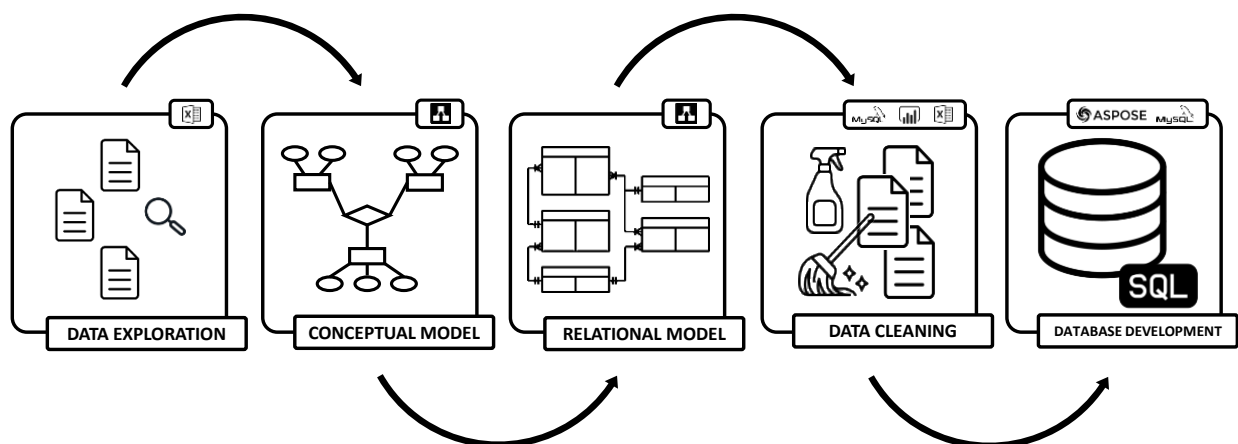


Figure 2. Workflow to create the database

## DATASET DESCRIPTION

All the following datasets have been downloaded from [www.italiandomani.gov.it/it/catalogo-open-data.html](http://www.italiandomani.gov.it/it/catalogo-open-data.html), and each of them is described in detail in a table.

In each table there are four sections, separated by a double line.

- In the first section there is a two-digit identification code (from 01 to 15) and the name of the dataset presented on the website.
- The second one contains the name of the .csv file, a brief description of the topic, the related keywords, and the last update date of the dataset.
- The third one contains the number of columns and the number of rows (excluding the first row where the attribute names are shown).
- The fourth one describes each attribute:
  - Name: name of the attribute in the .csv file.
  - Description: it is a description of the dataset that can contain additional information such as the domain of the attributes.
  - Type: can be 'Varchar' if the string has alphabetic or alphanumeric characters, while 'Num' if the string is a number.
  - NULL: it is the number of null values that are counted for that attribute (in the files they are often indicated as n.c. or N/A) and it is an index of completeness of the dataset (depending on the number of rows, the higher the number of null values, the more incomplete the dataset is).

Before starting with the description of all datasets, and as will become clearer later, it can be noted that there are some attributes that are common across many datasets. Therefore, for all these attributes, Table A shows their name, description, and type, consistently with what has been said previously.

**Table A.** Description of the most frequent attributes in datasets

Attribute Name	Description	Type
Programma	Unique code that identifies the belonging of the rows to the National Recovery and Resilience Plan submitted to the European Commission. <i>Domain = {PNRR}</i>	Varchar
Missione	Identification code of the PNRR Mission. <i>Domain = {M1; M2; M3; M4; M5; M6}</i>	Varchar
Descrizione Missione	A short description of the PNRR Mission. <i>Domain = {Digitalizzazione, innovazione, competitività e cultura; Rivoluzione verde e transizione ecologica; Infrastrutture per una mobilità sostenibile; Istruzione e ricerca; Inclusione e coesione; Salute}</i>	Varchar



Componente	Identification code of the PNRR Mission's Component.	Varchar
Descrizione Componente	Description of the PNRR Mission's Component.	Varchar
ID Misura	An Alphanumeric code that identifies in the System "Regis" the Measure intended as specific investments and/or reforms provided by the PNRR and realized through the implementation of interventions/ projects financed therein.	Varchar
Codice Univoco Misura	Alphanumeric code that uniquely identifies the Measures intended as specific investments and/or reforms.	Varchar
Descrizione Misura	Description of the PNRR Measure.	Varchar
ID submisura	Alphanumeric code that identifies in the System "Regis" the sub-measure, as a sub-area of intervention of the Measure.	Varchar
Codice CID	Alphanumeric code that identifies the sub-measure, according to the classification in the Annex to the Council Decision.	Varchar
Codice Univoco submisura	Alphanumeric code that uniquely identifies the sub-measure.	Varchar
Descrizione submisura	Description of the sub-measure.	Varchar
Tipo Livello	Articulation of the measure in sub-measures. If Type level is equal to 'Sub-Misura' the measure is divided into several Italian or European sub-measures. If Type Level is equal to 'Misura' the Measure is not articulated in sub-measures: that means that the last two characters of the Codice Univoco submisura are '00'. <i>Domain = {Sub-Misura; Misura}.</i>	Varchar
Importo Totale	Total amount in euro of the resources provided by the PNRR.	Num

**Table 1.** Monitoring of PNRR measures through sustainable development indicators (SDGs) and Agenda 2030

<b>01</b>	<b>MONITORING OF PNRR MEASURES THROUGH SUSTAINABLE DEVELOPMENT INDICATORS (SDGs) AND AGENDA 2030</b>			
<b>Dataset name</b>	Mappatura_PNRR_SNSvS_Agenda2030			
<b>Subjetc</b>	The dataset shows the details of the monitoring framework prepared together with the values of the statistical indicators of Welfare/Sustainability identified. The dataset shows the full overview of the links of the PNRR measures and submissions with the indicators of well-being and sustainability and their placement in the SDG framework of the 2030 agenda, the result of the ISTAT-MEF RGS collaboration.			
<b>Key words</b>	Monitoring, Agenda 2030, SDG			
<b>Last update date</b>	04 August 2022			
<b>Number of rows</b>	467			
<b>Number of columns</b>	31			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Programma	See Table A		0
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0

ID misura	See Table A		0
Codice Univoco Misura	See Table A		0
Descrizione Misura	See Table A		0
ID submisura	See Table A		0
Codice CID	See Table A		0
Codice Univoco submisura	See Table A		0
Descrizione submisura	See Table A		0
Tipo livello	See Table A		0
Importo totale	See Table A		371
Investimento/ Riforma	It specifies if the Submisura or Misura is an investment or a reform. <i>Domain = {Investimento; Riforma}.</i>	Varchar	0
Prestito/ Sovvenzione	It specifies if the financial support is a loan or grant. <i>Domain = {Prestito; sovvenzione, senza attribuzione}.</i>	Varchar	0
Amministrazione Titolare	Name of Administration who manages the Measure	Varchar	0
PREVALENTE	BES indicator indicated as more relevant to the purpose of the measure ('si' = more relevant Welfare/Sustainability indicator, 'no' = secondary indicator). <i>Domain = {si; no}</i>	Varchar	0
COD_INDICATORE	ISTAT Code of the Wellness/Sustainability (BES) Indicator	Varchar	97
DES_INDICATORE	Description of BES Indicator	Varchar	0
DES_INDAGINE	Istat source description of BES indicator	Varchar	106
DEFINIZIONE	Extension definition of BES indicator	Varchar	97
NOTA	Additional information about BES indicator	Varchar	264
DES_FONTE	Description of ISTAT source	Varchar	97
UNITA	Unit of measure of BES indicator	Varchar	97
INDICATOR	SDG code Indicator	Varchar	134
DES_INDICATOR	Description of SDG code Indicator	Varchar	135
TARGET	Target code of SDG indicator	Varchar	134
DES_TARGET	Description of SDG Target	Varchar	135
GOAL	Goal code of SDG Target	Varchar	0
DES_GOAL	Description of SDG Goal	Varchar	0

In this dataset there is a problem of inconsistency, since for several tuples, despite Prevalente = 'si', which implies the existence of a BES indicator, in COD\_Indicator there is no code and there is however a description. As we will see in the following chapters, for our Conceptual Model we discard all columns related to Bes Indicator because it is not relevant for our analysis.

Table 2. Milestone and Target PNRR programming

02 MILESTONE AND TARGET PNRR PROGRAMMING				
Dataset name	PNRR_MeT_OA_programmazione_v02			
Subjetc	It shows for each sub-measure the related Milestone and Target associated. Milestone is a quality indicator while Target is a quantity indicator. Each indicator describes the state of achievement and result for every Reform or Investment in PNRR.			
Key words	Enactment, Milestone, Target			
Last update date	03 August 2022			
Number of rows	1489			
Number of columns	24			
Attributes	Name	Description	Type	NULL
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	Codice Univoco Misura	See Table A		0
	ID misura	See Table A		0
	Descrizione Misura	See Table A		0
	Codice Univoco Submisura	See Table A		0
	ID Submisura	See Table A		0
	Descrizione Submisura	See Table A		0
	Amministrazione Titolare	See Table 01		0
	Codice univoco Milestone / Target	Unique code for identifying the goal.	Varchar	0
	Milestone / Target	It specifies if the goal is a Milestone or a Target. Domain = {Milestone; Target}	Varchar	0
	Nome Milestone/ Target	Name of Milestone/Target	Varchar	0
	Descrizione specifica di ogni Milestone e Target	Description of Milestone/Target	Varchar	0
	ITA/UE	It specifies if the Milestone/Target is ITA or EU. Domain = {ITA; UE}.	Varchar	0
	Indicatori qualitativi Milestone	Quality indicator description of Milestone	Varchar	856
	Unità di misura Target (riclassificata)	Unit measure of Target quantity indicator	Varchar	0

	Riferimento di partenza Target	Starting point of the Target	Num	0
	Valore-obiettivo Target	Goal value of the Target quantity indicator	Num	641
	Trimestre di conseguimento	Quarter of achievement	Varchar	0
	Anno di conseguimento	Year of achievement	Num	0
	OA - Meccanismo di verifica	Description of the periodic verification mechanisms for achieving Milestone and Target	Varchar	789
	OA - Ulteriori Informazioni	Additional information related to Milestone and Target periodic verification mechanisms	Varchar	1209

**Table 3.** Milestone and Target implementation at T4/2021

<b>03 MILESTONE AND TARGET IMPLEMENTATION AT T4/2021</b>				
<b>Dataset name</b>	MeT_Attuazione_2021-T4			
<b>Subjetc</b>	The dataset monitors the UE Milestone and Target at the Fourth Quarter of 2022 explaining the State of achievement and the main evidence.			
<b>Key words</b>	Enactment, Milestone, Target			
<b>Last update date</b>	22 July 2022			
<b>Number of rows</b>	51			
<b>Number of columns</b>	22			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	ID misura	See Table A		0
	Codice Univoco Misura	See Table A		0
	Descrizione Misura	See Table A		0
	Amministrazione Titolare	See Table 01		0
	Codice univoco Milestone / Target	See Table 02		0
	Milestone / Target	See Table 02		0
	Nome Milestone/ Target	See Table 02		0
	Avanzamento	State of achievement of Milestone/Target	Varchar	0
	Principali evidenze	Information about main evidence from that Milestone/Target	Varchar	0

	Allegati, from 1 to 25	Some web articles that prove the achievement of that Milestone/Target	Varchar	/
--	------------------------	---	---------	---

**Table 4.** Milestone and Target implementation at T2/2022

<b>04 MILESTONE AND TARGET IMPLEMENTATION AT T2/2022</b>				
<b>Dataset name</b>	MeT_Atтуazione_2022-T2			
<b>Subjetc</b>	The dataset monitors the UE Milestone and Target at the Second Quarter of 2022 explaining the State of achievement and the main evidence.			
<b>Key words</b>	Enactment, Milestone, Target			
<b>Last update date</b>	08 July 2022			
<b>Number of rows</b>	45			
<b>Number of columns</b>	38			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	ID misura	See Table A		0
	Codice Univoco Misura	See Table A		0
	Descrizione Misura	See Table A		0
	Amministrazione Titolare	See Table 01		0
	Codice univoco Milestone / Target	See Table 02		0
	Milestone / Target	See Table 02		0
	Nome Milestone/ Target	See Table 02		0
	Avanzamento	See Table 03		0
	Principali evidenze	See Table 03		0
	Allegati, from 1 to 9	See Table 03		/

**Table 5.** PNRR Financial Framework

<b>05 PNRR FINANCIAL FRAMEWORK</b>	
<b>Dataset name</b>	OpenData_PNRR_QuadroFin_Programmazione_v02
<b>Subjetc</b>	This dataset explains where the money comes from, for that submission, and how they are structured.

<b>Key words</b>	Enactment, Financial framework			
<b>Last update date</b>	26 April 2022			
<b>Number of rows</b>	285			
<b>Number of columns</b>	20			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Programma	See Table A		0
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	ID misura	See Table A		0
	Codice Univoco Misura	See Table A		0
	Descrizione Misura	See Table A		0
	ID submisura	See Table A		0
	Codice CID	See Table A		0
	Codice Univoco submisura	See Table A		0
	Descrizione submisura	See Table A		0
	Tipo livello	See Table A		0
	Investimento/ Riforma	See Table A		0
	Prestito/ Sovvenzione	See Table A		0
	Amministrazione Titolare	See Table A		0
	Importo Totale	See Table A		0
	Importo Progetti in essere	Amount in euro of resources allocated to projects financed by national expenditure authorizations	Num	247
	Importo FSC	part of the money that comes from 'Fondo Sviluppo e Coesione 2021-2027'	Num	262
	Importo Progetti nuovi	Amount in euro of resources allocated to projects financed by national expenditure authorisations	Num	73

Table 6. Interoperability - Administration Owner

<b>06</b>	<b>INTEROPERABILITY - ADMINISTRATION OWNER</b>			
<b>Dataset name</b>	AmmTitolari_interoperabilitàBDAP			
<b>Subject</b>	This dataset gives all the information about the entity who control the PNRR Measures.			
<b>Key words</b>	Enactment			
<b>Last update date</b>	26 April 2022			
<b>Number of rows</b>	25			
<b>Number of columns</b>	9			

Attributes	Name	Description	Type	NULL
	Amministrazione Titolare	See Table 01		0
	Id_Ente BDAP	'Banca Dati delle Amministrazioni Pubbliche BDAP' Code to identify the Administration	Varchar	0
	Denominazione per esteso	Name of Administration in BDAP system	Varchar	0
	Denominazione BDAP	Administration name according to BDAP system	Varchar	0
	CF	Tax Code of the Administration	Varchar	0
	Codice_Ente_IPA	IPA code of the Administration	Varchar	0
	Codice_UO	Personal data of parties which composed the Administration	Varchar	16
	Denominazione Unità Organizzativa	IPA Name of Administration		16
	Note	Others information	Varchar	23

Table 7. TAG for PNRR climate and Digital Support

07 TAG FOR PNRR CLIMATE AND DIGITAL SUPPORT				
Dataset name	TAG_sostegno_clima_digitale_v02			
Subjetc	This Dataset tell us how ad if a sub-measure intervenes in Climate field or Digital field. There is also the respectively total amount in each field, calculated multiplying 'Importo totale' by the related Coefficient			
Key words	TAG, Climate, Digital			
Last update date	27 April 2022			
Number of rows	281			
Number of columns	22			
Attributes	Name	Description	Type	NULL
	Programma	See Table A		0
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	ID misura	See Table A		0
	Codice Univoco Misura	See Table A		0
	Descrizione Misura	See Table A		0
	ID submisura	See Table A		0
	Codice CID	See Table A		0
	Codice Univoco submisura	See Table A		0

	Descrizione submisura	See Table A		0
	Amministrazione titolare	See Table A		0
	Periodo di tempo: inizio	Start date of the implementation period of the measure or submission		0
	Periodo di tempo: fine	End date of the implementation period of the measure or submission		0
	Importo totale	See Table A		0
	Campo d'intervento Clima	Field of action in the Climate	Char	0
	Coefficiente TAG Clima	It specifies, as a percentage, the part of the total amount that intervenes in the climate field	Num	173
	Campo d'intervento Digitale	Field of action in the Digital	Varchar	189
	Coefficiente TAG digitale	It specifies, as a percentage, the part of the total amount that intervenes in the digital field	Num	189
	Importo Contribuzione al Clima	Total amount of euro resources of the measure/sub-measure contributing effectively to climate objectives	Num	182
	Importo contribuzione al Digitale	Total amount of euro resources of the measure/sub-measure that effectively contributes to the digital transition	Num	194

Here were found inconsistency errors because despite having a climate tag coefficient, there was no corresponding part of climate contribution amount.

**Table 8.** Localization of PNRR Projects

<b>08 LOCALIZATION OF PNRR PROJECTS</b>				
<b>Dataset name</b>	PNRR_Localizzazione			
<b>Subjetc</b>	<p>The dataset contains information on the location of the project: it indicates Region, Province, Address and Zip Code.</p> <p>The relationship between CUP (that is the identifying code of the project) and territories is not one-to-one, because several projects could be in the same territory.</p>			
<b>Key words</b>	Project, Location			
<b>Last update date</b>	13 May 2022			
<b>Number of rows</b>	5246			
<b>Number of columns</b>	10			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Codice Univoco submisura	See Table A		0
	CUP	A 15 digits code assigned to each public investment project by the CUP System, managed	Varchar	0



		by the Department for Planning and Coordination of Economic Planning (DIPE)		
	Regione	ISTAT code (2 digits) of the region	Num	0
	Descrizione Regione	Name of the region	Varchar	0
	Provincia	ISTAT code (3 digits) of the province	Num	0
	Descrizione Provincia	Name of the province	Varchar	0
	Comune	ISTAT code (3 digits) of the municipal territory	Num	0
	Descrizione Comune	Name of the town	Varchar	0
	Indirizzo	Address	Varchar	5246
	Cap	ZIP code (5 digits)	Num	0

In this dataset there is a problem related to completeness, because 100% of the “Indirizzo” values are null, while there are no consistency problems because all the Cap are composed by five digits and are consistent with the town. However, there is a peculiarity because if Region is equal to ‘000’ the project has a national scope, if Provincia is equal to ‘000’ the project is located in all the provinces and finally when Comune is equal to ‘000’ the project is located in all the towns. In all this cases the Cap is equal to ‘00000’, which is a cap that does not exist in Italy.

Table 9. PNRR Projects

09 PNRR PROJECTS				
Dataset name	PNRR_Progetti			
Subjetc	The dataset associates the projects to each measure or submission and to their informative kit of the Plan. For each CUP the dataset reported nature, title and total funding, broken down by source			
Key words	Project, Enactment			
Last update date	13 May 2022			
Number of rows	5246			
Number of columns	28			
Attributes	Name	Description	Type	NULL
	Programma	See Table A		0
	Missione	See Table A		0
	Descrizione Missione	See Table A		0
	Componente	See Table A		0
	Descrizione Componente	See Table A		0
	ID misura	See Table A		0
	Codice Univoco Misura	See Table A		0
	Descrizione Misura	See Table A		0
	ID submisura	See Table A		0

	Codice CID	See Table A		0
	Codice Univoco submisura	See Table A		0
	Descrizione submisura	See Table A		0
	Amministrazione Titolare	See Table 01		0
	CUP	See Table 08		0
	CUP Codice Natura	Identification code of the nature of the project	Num	0
	CUP Descrizione Natura	Description of the nature of the project	Varchar	0
	Codice Locale Progetto	Project identification code in the monitoring system	Varchar	0
	Titolo Progetto	Project name	Varchar	0
	Finanziamento UE	Amount financed from EU sources (not PNRR)	Num	5246
	Finanziamento Regione	Amount financed by the region	Num	5246
	Finanziamento Provincia	Amount financed by the province	Num	5246
	Finanziamento Comune	Amount financed by the town	Num	5246
	Finanziamento Altro Pubblico	Amount financed from other public source	Num	5246
	Finanziamento Privato	Amount financed from private sources	Num	5246
	Finanz da Reperire	Amounts to be recovered. It represents the part of the project cost not yet covered by funding.	Num	5246
	Finanziamento PNRR	Amount financed by the PNRR	Num	0
	Finanziamento PNC	Amount financed by PNC	Num	5246
	Finanziamento Totale	Total amount of funding	Num	0

In this dataset there is a problem of completeness and accuracy because apart from Finanziamento PNRR, all funding has 100% null values: in this case it is not clear whether NULL should be understood as 0€ or it is a missing information. In any case Finanziamento PNRR and Finanziamento Totale (that is the sum of all the previous funding) always have the same amount, so this suggests that the first deduction is correct.

Table 10. Subject of the PNRR

10	SUBJECTS OF THE PNRR			
<b>Dataset name</b>	PNRR_Soggetti			
<b>Subjetc</b>	The dataset associates programmer, actuator and beneficiary to each project. For each of them, information on the Tax Code, name, legal form, and ATECO code of the economic activity is reported			
<b>Key words</b>	Project, Stakeholders			
<b>Last update date</b>	13 May 2022			
<b>Number of rows</b>	5246			
<b>Number of columns</b>	17			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Codice Univoco submisura	See Table A		0
	CUP	See Table 08		0
	Codice Fiscale Programmatore	Programmer Tax Code (11 digits)	Num	0
	Denominazione Programmatore	Programmer name	Varchar	0
	Forma Giuridica Programmatore	Code identifying the legal form of the programmer (2 digits)	Num	0
	Descrizione Forma Giuridica Programmatore	Description of the legal form of the programmer	Varchar	0
	Codice ATECO Programmatore	Ateco 2007 Code of the economic activity of the programmer (8 digits)	Num	0
	Codice Fiscale Attuatore	Actuator Tax Code	Num / Varchar	0
	Denominazione Attuatore	Actuator name	Varchar	0
	Forma Giuridica Attuatore	Code identifying the legal form of the actuator (2 digits)	Num	0
	Descrizione Forma Giuridica Attuatore	Description of the legal form of the actuator	Varchar	0
	Codice ATECO Attuatore	Ateco 2007 Code of the economic activity of the actuator (8 digits)	Num	0
	Codice Fiscale Beneficiario	Beneficiary Tax Code	Num / Varchar	22
	Denominazione Beneficiario	Beneficiary name	Varchar	22
	Forma Giuridica Beneficiario	Code identifying the legal form of the beneficiary (2 digits)	Num	26
	Descrizione Forma Giuridica Beneficiario	Description of the legal form of the beneficiary	Varchar	26
	Codice ATECO Beneficiario	Ateco 2007 Code of the economic activity of the beneficiary (8 digits)	Num	22

Table 11. PNRR Competitions

11 PNRR COMPETITIONS				
Dataset name	PNRR_Gare			
Subjetc	The dataset associates projects with competitions. For each competition, the Tax Code or VAT number, the name, the legal form, the ATECO sector of the developer or the successful tenderer, the description of the award procedure and the awarded amount are reported			
Key words	Competition, Enactment			
Last update date	13 May 2022			
Number of rows	4			
Number of columns	11			
Attributes	Name	Description	Type	NULL
	Codice Univoco submisura	See Table A		0
	CUP	See Table 08		0
	Codice Fiscale Realizzatore (o Aggiudicatario)	Tax Code of the subject and/or economic operator involved in the realization of the project	Num	0
	Denominazione Realizzatore (o Aggiudicatario)	Name of the subject and/or economic operator involved in the realization of the project	Varchar	0
	Forma Giuridica Realizzatore (o Aggiudicatario)	Legal form code of the entity and/or economic operator involved in the realization of the project	Num	0
	Descrizione Forma Giuridica Realizzatore (o Aggiudicatario)	Description of the legal form code of the entity and/or economic operator involved in the realization of the project	Varchar	0
	Codice ATECO Realizzatore (o Aggiudicatario)	Economic activity ATECO code of the entity and/or economic operator involved in the realization of the project	Varchar	0
	CIG	Identification code of the Competition	Varchar	0
	Descrizione Procedura di Aggiudicazione	Description of the award procedure	Varchar	0
	Importo Aggiudicato	Amount awarded at the end of the competition evaluation process	Num	4
	Codice Fiscale / Partita Iva Aggiudicatario	Tax number or VAT number of the successful bidder	Num	0

This dataset contains limited information because it shows only four lines related to a single project (CUP = 'J51B21002940006')

Table 12. TAG Projects PNRR

12	TAG PROJECTS PNRR			
<b>Dataset name</b>	PNRR_TAG_Progetti			
<b>Subjetc</b>	The dataset associates each project, with the digital TAG, the support coefficient, the field of intervention and the amount of resources associated			
<b>Key words</b>	Project, TAG, Enactment			
<b>Last update date</b>	13 May 2022			
<b>Number of rows</b>	5225			
<b>Number of columns</b>	8			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Codice Univoco submisura	See Table A		0
	CUP	See Table 08		0
	Tipo TAG Clima	String ("CLIMA TAG") that indicates if the project acts on the climate	/	5225
	Campo d'intervento Clima	Field of action for climate marking	/	5225
	Coefficiente TAG Clima	Percentage coefficient for calculating support for environmental objectives	/	5225
	Tipo Tag Digitale	String ("DIGITAL TAG") that indicates if the project acts on the climate	Varchar	0
	Campo d'intervento Digitale	Field of action for digital marking	Varchar	0
	Coefficiente TAG Digitale	Percentage coefficient for calculating support for digital objectives	Num	0

Although the dataset contains attributes relating to the Digital TAG and the Climate one, for the latter, all values are null, so the dataset provides information only on projects with a Digital TAG. The attribute "Coefficiente TAG Digitale" indicates the percentage of the total project amount that contributes to the digital goals.

Table 13. PNC Dataset - Objectives

13	PNC DATASET - OBJECTIVES			
<b>Dataset name</b>	Dataset PNC - Obiettivi			
<b>Subjetc</b>	The Italian Government integrates and enhances the contents of the PNRR through the Complementary National Plan (PNC) by allocating an additional 30.6 billion of national resources. The dataset associates the programs foreseen by the PNC with the objectives to be achieved, specifying trimester and year			
<b>Key words</b>	PNC			
<b>Last update date</b>	22 December 2021			
<b>Number of rows</b>	308			
<b>Number of columns</b>	10			

Attributes	Name	Description	Type	NULL
	Codice Univoco programma	Alphanumeric code that uniquely identifies the programs provided for in the PNC	Varchar	0
	Codice univoco investimento   Intervento	Identification code at the investment (and intervention, where applicable) level foreseen in the PNC scope	Varchar	0
	Amministrazione Titolare	Ministries and structures of the Presidency of the Council of Ministers responsible for the implementation of the Measures	Varchar	0
	Programma	Name of the program	Varchar	0
	Investimento	Name of the investment contained in the program	Varchar	155
	Intervento	Name of the intervention contained in the investment scope	Varchar	248
	Obiettivi	Description of the initial, intermediate, and final objectives identified by the Titular Administrations. If the program is co-financed, the objective indicates the enhanced PNRR Measure	Varchar	0
	Trimestre	2 digits code the indicates the trimester in which to reach the target	Varchar	6
	Anno	Year associated with the trimester within which the objective must be achieved	Num	6
	Codice Univoco obiettivo	Alphanumeric code uniquely identifying the objective associated with the program/investment/intervention of the PNC in the overall financial framework	Varchar	0

Table 14. PNC Dataset – Programs description

14	PNC DATASET – PROGRAMS DESCRIPTION			
Dataset name		OpenData_PNC_DescrizioneProgrammi		
Subject		The dataset contains the content and purposes of the PNC programs		
Key words		PNC		
Last update date		22 December 2021		
Number of rows		30		
Number of columns		3		
Attributes	Name	Description	Type	NULL
	Codice Univoco programma	See Table 13		0
	Programma	See Table 13		0
	Descrizione Programma	Brief description of the program and the main fields of action	Varchar	0


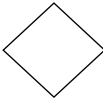
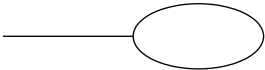
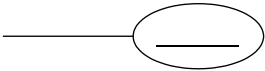
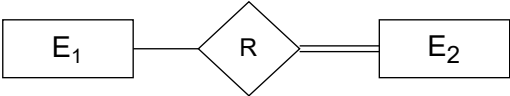
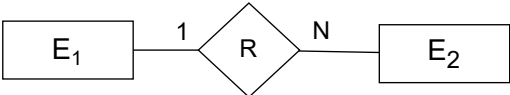
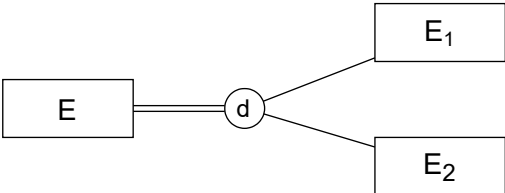
Table 15. PNC Dataset - Financial framework

15	PNC DATASET - FINANCIAL FRAMEWORK			
<b>Dataset name</b>	OpenData_PNC_QuadroFinanziario			
<b>Subjetc</b>	The plan provides for 30 programs, 24 of which are financed exclusively by the PNC (and therefore from the State budget) and 6 programs are co-financed with the PNRR (i.e., already provided for in the PNRR and for which the PNC provides for additional resources) (www.italiadomani.gov.it). The dataset shows all the programs contained in the PNC, highlighting: the type of measure, their placement related to the PNRR, the Administration and the total amount			
<b>Key words</b>	PNC			
<b>Last update date</b>	22 December 2021			
<b>Number of rows</b>	30			
<b>Number of columns</b>	7			
<b>Attributes</b>	<b>Name</b>	<b>Description</b>	<b>Type</b>	<b>NULL</b>
	Riforma / Investimento	Type of measure. <i>Domain = {Investment; Reform}</i>	Varchar	0
	Codice Univoco programma	<i>See Table 13</i>		0
	Programma esclusivo / cofinanziato	Alphanumeric code that identifies the location of the program. The variable takes two values: "PNC-PNRR" if the program is co-financed and therefore there is an integration of the resources in the intervention already foreseen in the PNRR; "PNC-DM MEF 15 Luglio 2021" if it is a further action than those already contained in the PNRR. <i>Domain = {PNC-DM MEF 15 Luglio 2021; PNC-PNRR}</i>	Varchar	0
	Programma	Name of the program	Varchar	0
	Amministrazione Titolare	Ministries and structures of the Presidency of the Council of Ministers responsible for the implementation of the Measures	Varchar	0
	Importo	Total amount of PNC resources per program (million euro)	Num	0
	Componente PNRR	<i>See Table A</i>	Varchar	24

In this dataset a peculiarity lies in the fact that the first column contains 100% of the values equal to 'Investimento', therefore all programs represented an investment.

## THE ENHANCED ENTITY-RELATIONSHIP MODEL

Starting from the description of the tables analyzed in the previous chapter we have developed the Enhanced Entity-Relationship model. We created different entities, associating to each of them a series of attributes that characterized them based on the information collected through the data description. To do this we chose the most significant columns and discarded the least relevant ones, either because they were full of null values or because they were not useful for the decision-making process. Figure 3 shows the summary of the notation for the EER Diagrams that we used, Figure 4 shows the digraph that we elaborated to develop the PNRR database later, while all the Tables that follow the diagram (from 16 to 22), show the attribute indicated in the EER and the dataset column, described in the previous chapter, to which it refers.

SYMBOL	MEANING
	Entity
	Relationship
	Attribute
	Key Attribute
	Total Participation of $E_2$ in $R$
	Cardinality Ratio 1:N for $E_1 : E_2$ in $R$
	Entity $E$ specializes itself in $E_1$ and $E_2$

**Figure 3.** Summary of the notation for EER diagrams



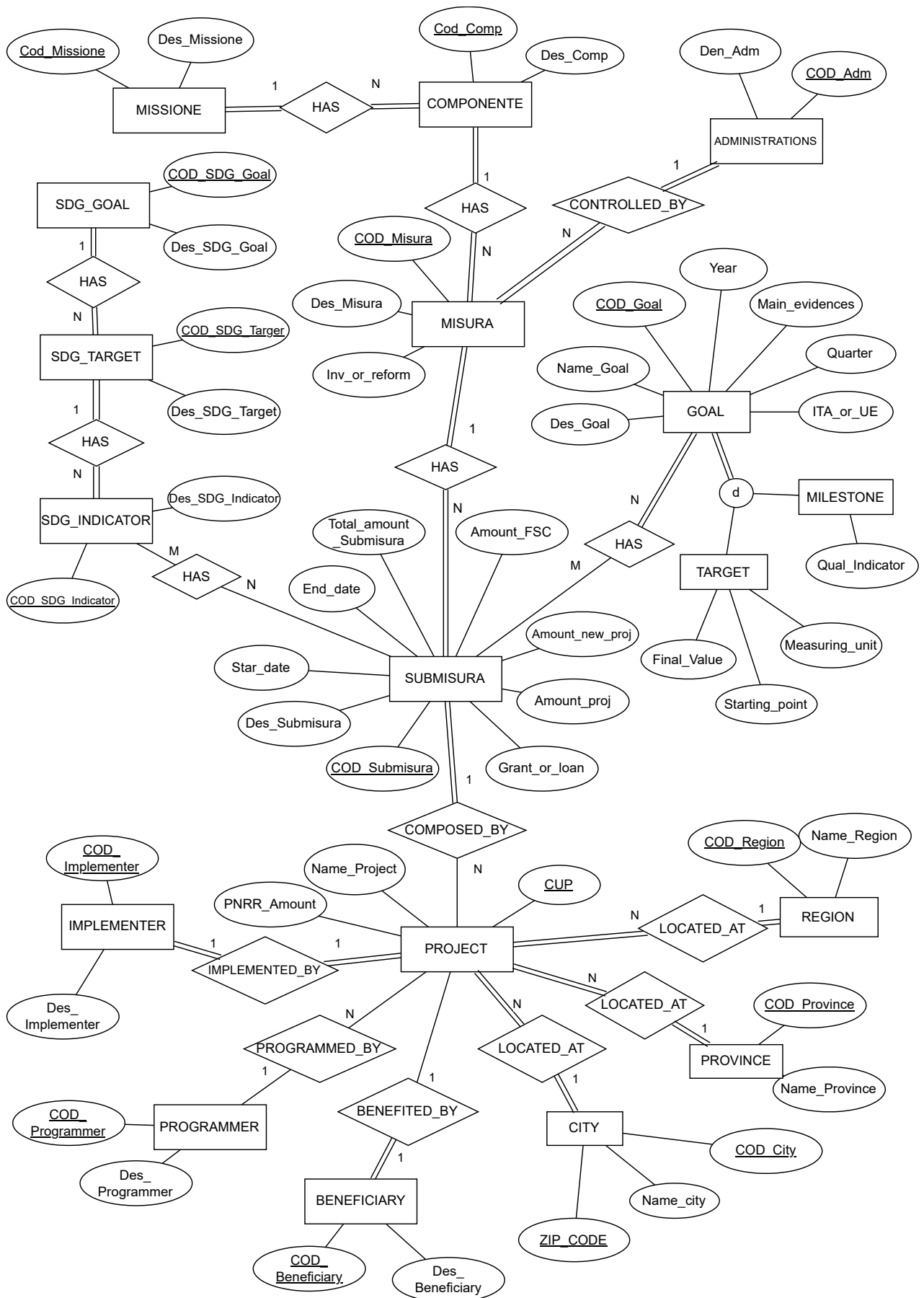


Figure 4. PNRR EER Model

**Table 16.** Relationship between EER attributes and datasets columns for Missione, Componente, Misura and Submisura

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
<b>MISSIONE</b>	COD_Missione	Missione	A
	Des_Missione	Descrizione Missione	A
<b>COMPONENTE</b>	COD_Comp	Componente	A
	Des_Comp	Descrizione Componente	A
<b>MISURA</b>	COD_Misura	Codice univoco misura	A
	Des_Misura	Descrizione misura	A
	Inv_or_reform	Investimento / Riforma	01
<b>SUBMISURA</b>	COD_Submisura	Codice Univoco Submisura	A
	Des_Submisura	Descrizione Submisura	A
	Grant_or_loan	Sovvenzione / Prestito	01
	Total_amount_submisura	Importo Totale	A
	Amount_proj	Importo progetti in essere	05
	Amount_FSC	Importo FSC	05
	Amount_new_proj	Importo Progetti nuovi	05
	Start_date	Periodo di tempo: inizio	06
	End_date	Periodo di tempo: fine	06

**Table 17.** Relationship between EER attributes and datasets columns for Goal, Milestone, Target

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
<b>GOAL</b>	COD_Goal	Codice univoco Milestone / Target	02
	Name_Goal	Nome Milestone / Target	02
	Des_Goal	Descrizione specifica di ogni Milestone e Target	02
	ITA_or_UE	ITA / UE	02
	Main_evidences	Principali evidenze	03
	Quarter	Trimestre	02
	Year	Anno di conseguimento	02
<b>MILESTONE</b>	Qual_indicator	Indicatore qualitativo Milestone	02
<b>TARGET</b>	Measuring_unit	Unità di misura Target (riclassificata)	02
	Starting_point	Riferimento di partenza Target	02
	Final_Value	Valore-obiettivo Target	02

**Table 18.** Relationship between EER attributes and datasets columns for Administrations

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
<b>ADMINISTRATIONS</b>	COD_Adm	Amministrazione Titolare	07
	Den_Adm	Denominazione per esteso	07

**Table 19.** Relationship between EER attributes and datasets columns for SDG Goal, SDG Target and SDG Indicator

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
SDG_GOAL	COD_SDG_Goal	GOAL	01
	Des_SDG_Goal	Des_Goal	01
SDG_TARGET	COD_SDG_Target	TARGET	01
	Des_SDG_Target	Des_Target	01
SDG_INDICATOR	COD_SDG_Indicator	INDICATOR	01
	Des_SDG_Indicator	Des_indicator	01

**Table 20.** Relationship between EER attributes and datasets columns for Project

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
PROJECT	CUP	CUP	09
	Name_Project	Titolo Progetto	09
	PNRR_Amount	Finanziamento PNRR	09

**Table 21.** Relationship between EER attributes and datasets columns for Region, Province, and City

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
REGION	COD_Region	Regione	08
	Name_Region	Descrizione Regione	08
PROVINCE	COD_Province	Provincia	08
	Name_Province	Descrizione Provincia	08
CITY	COD_City	Comune	08
	ZIP_CODE	CAP	08
	Name_City	Descrizione comune	08

**Table 22.** Relationship between EER attributes and datasets columns for Implementer, Programmer and Beneficiary

Entity	EER-Model Attribute Name	Dataset Attribute Name	Table ID
IMPLEMENTER	COD_Implementer	Codice Fiscale Attuatore	10
	Des_Implementer	Denominazione Attuatore	10
PROGRAMMER	COD_Programmer	Codice Fiscale Programmatore	10
	Des_Programmer	Denominazione Programmatore	10
BENEFICIARY	COD_Beneficiary	Codice Fiscale Beneficiario	10
	Des_Beneficiary	Denominazione Beneficiario	10

# THE RELATIONAL MODEL

Starting from the EER model we developed the relational model following the steps described in the introductory chapter (Focus on the context: PNRR and databases). By mapping the entity types, we obtained the entity relation and then we reported the simple attributes as attributes of the relation and the key attribute became the primary key of the relation, distinguishable from the others being underlined. Following the rules described above and shown below, Figures 5 and 6 show the two steps of mapping: Figure 5 shows the mapping of Entity Relations and Figure 6 is the complete result of mapping the PNRR EER Model into a Relational Database Schema, including the Relationship Relation (M:N) and all the foreign keys. We decided to map the foreign keys in 1:1 relationships to the side where there was total participation and in 1:N relationships we mapped them, as usual, on N-side. Each arrow is oriented in such a way as to start from the external key and get to its primary key: This means that each tuple at the foreign key has the same domain as the primary key and is constrained to have the same values as the primary key or at most the NULL value (Referential integrity constraint).

1. Each entity becomes a table.
2. Entity attributes become table columns.
3. Columns inherit attributes characteristics.
4. The key attribute of the entity becomes the primary key of the table.
5. If the association is 1 to N, on the N side a column is added, corresponding to the primary key of side 1. This column is the foreign key of the relation.
6. If the association is 1 to 1, you can choose where to add the column.
7. If the association is M to N, a third table is added, which contains the keys of the other two tables

## MISSIONE

<u>COD_Missione</u>	Des_Missione
---------------------	--------------

## COMPONENTE

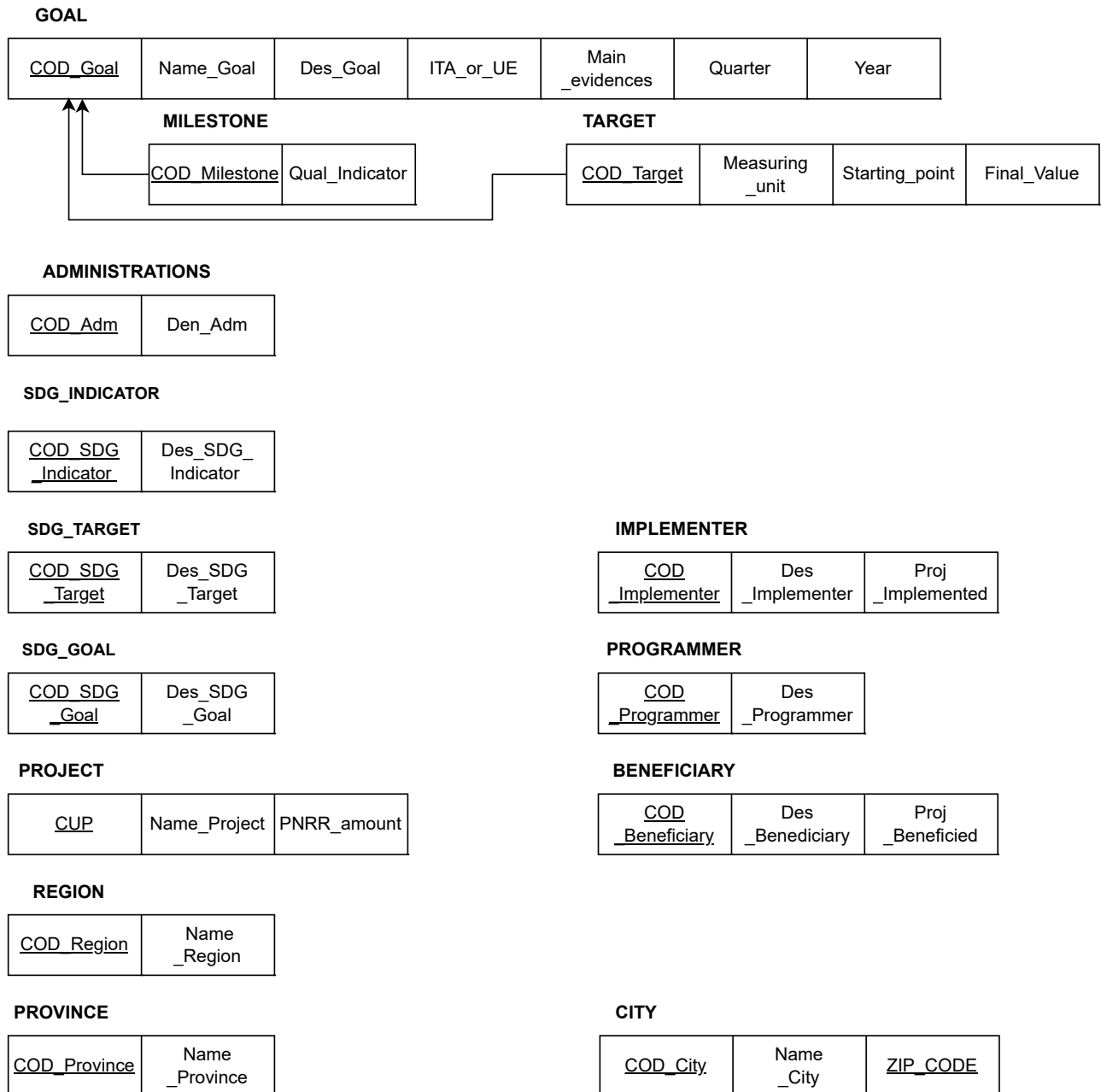
<u>COD_Comp</u>	Des_Comp
-----------------	----------

## MISURA

<u>COD_Misura</u>	Des_Misura	Inv_or_reform
-------------------	------------	---------------

## SUBMISURA

<u>COD_Submisura</u>	Des_Submisura	Grant_or_loan	Total_amount_Submeasure	Amount_proj	Amount_FSC	Amount_new_proj
					Start_date	End_date



**Figure 5.** Mapping of the entity type into entity relation

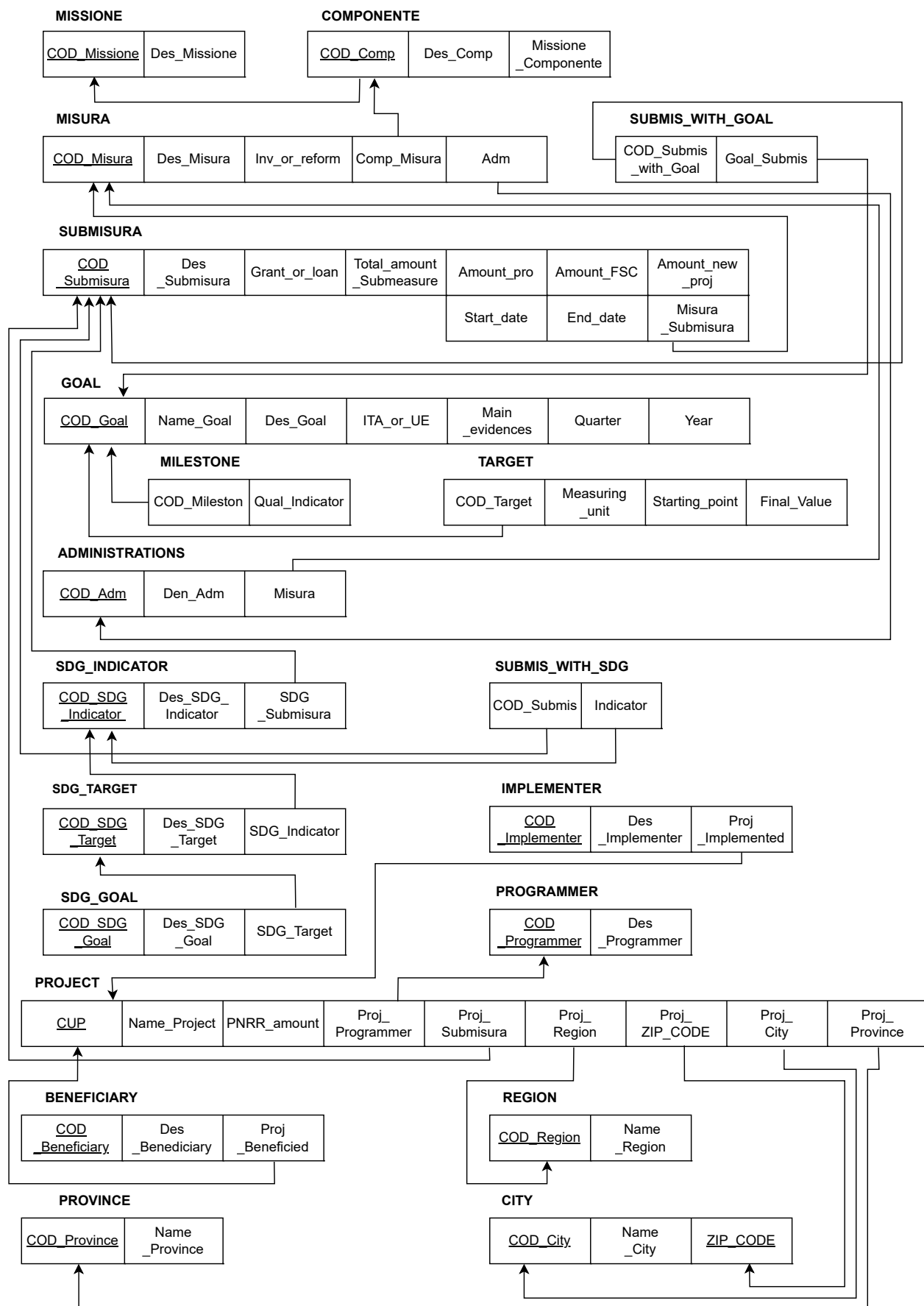


Figure 6. PNRR Relational Model

## DATABASE DEVELOPMENT

To develop the database we used MySQL, a relational database management system, and MySQL Workbench, a visual database design tool, which integrates SQL development, management, data modeling, creation, and maintenance of MySQL databases within a single synergistic environment. The first step, therefore, was the creation of the schema, through the script shown below.

```
CREATE SCHEMA pnrr_contest;
```

After that, we started creating all the tables inside the schema, according to the nomenclature and structure shown in the relational model in figure 6, and indicating the primary key, any foreign keys and the data types for each attribute. The creation of the Measure and Sub-measure tables is shown below as an example.

```
CREATE TABLE Misura (  
    COD_Misura CHAR(9) PRIMARY KEY,  
    Des_Misura VARCHAR (450),  
    Inv_or_Reform VARCHAR(13),  
    Comp_Misura CHAR (4),  
    FOREIGN KEY (Comp_Misura) REFERENCES Componente (COD_Comp)  
);
```

```
CREATE TABLE Submisura (  
    COD_Submisura CHAR (12) PRIMARY KEY,  
    Des_Submisura VARCHAR (700),  
    Grant_or_loan VARCHAR(20),  
    Amount_proj DOUBLE,  
    Amount_new_proj DOUBLE,  
    Amount_FSC DOUBLE,  
    Total_amount_submisura DOUBLE,  
    Start_date DATE,  
    End_date DATE,
```

```
Misura_Submisura CHAR(10),  
FOREIGN KEY (Misura_Submisura) REFERENCES Misura (COD_Misura)  
);
```

After creating all tables and placing the referential integrity constraints, the result is represented in figure 7.



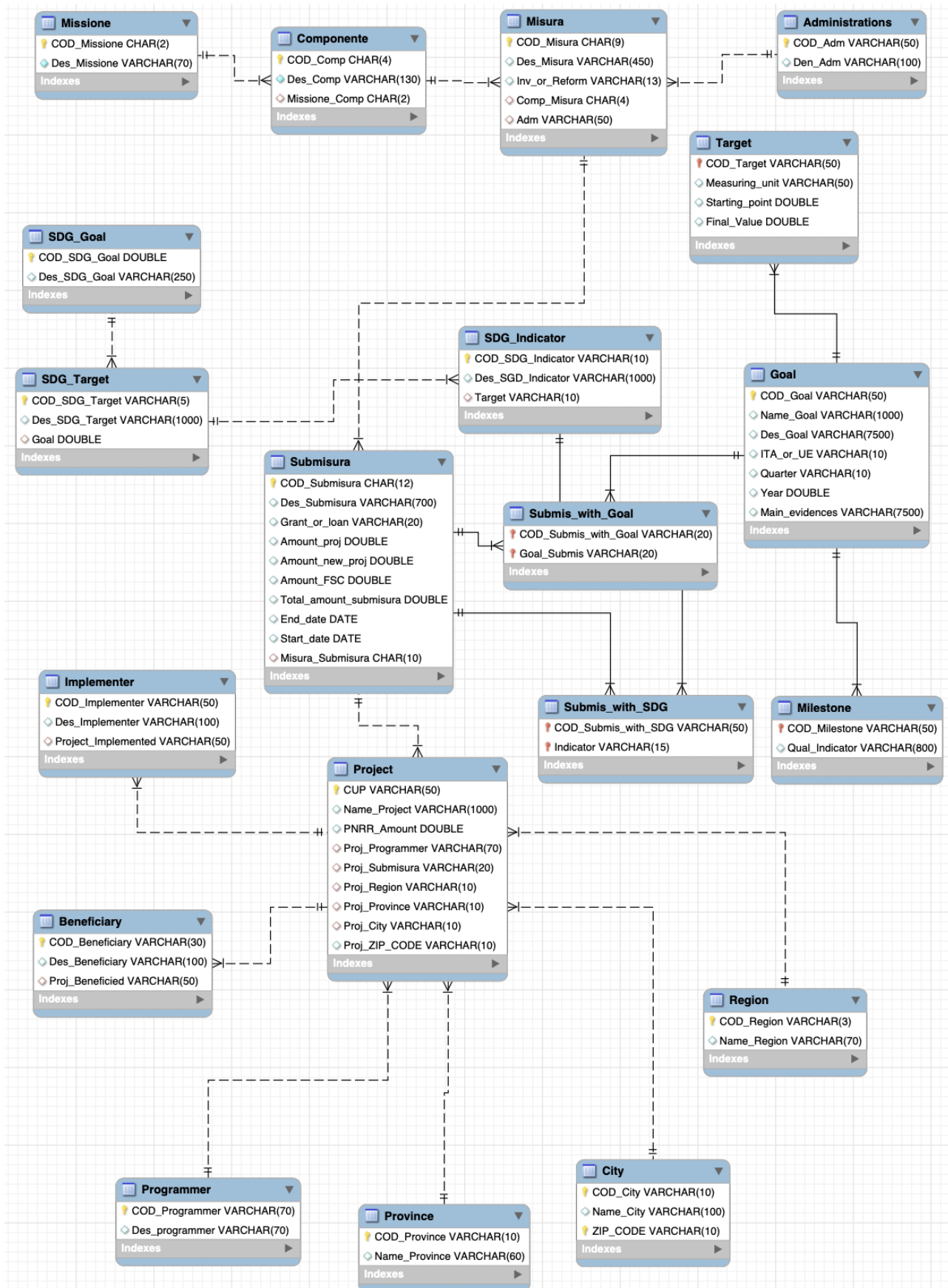


Figure 7. EER Diagram obtained using the MySQL Workbench Revers Engineer function

## Requested queries

To query the database that we created, we analyzed each relation itself and in relationship with the others. First of all, to do this we counted the number of rows for each table (retrieving it as 'Rows\_number\_[Name\_Table]') in fact, query 1 shows, as an example, the script we used for the Misura Table, while Table 23 returns the results of the same query applied to all tables.

### Query 1.

```
SELECT COUNT(*) AS 'Rows_number_Misura'  
  
FROM Misura;
```

**Table 23.** Query 1 results applied to all database tables

Subject	Number
Rows_number_Administrations	25
Rows_number_Beneficiary	5224
Rows_number_City	1760
Rows_number_Componente	16
Rows_number_Goal	1150
Rows_number_Implementer	5246
Rows_number_Milestone	482
Rows_number_Missione	6
Rows_number_Misura	191
Rows_number_Programmer	3
Rows_number_Project	5246
Rows_number_Province	108
Rows_number_Region	21
Rows_number_SDG_Goal	17
Rows_number_SDG_Indicator	242
Rows_number_SDG_Target	169
Rows_number_Submis_with_Goal	1489
Rows_number_Submis_with_SDG	268
Rows_number_Submisura	285
Rows_number_Target	668

As regards the 1:N relationship between regions and projects, total on both sides, the number of distinct regions participating in the relationship can be calculated through Query 2 which returns the value of 21.

#### Query 2.

```
SELECT COUNT(DISTINCT Proj_Region)

FROM Project;
```

Therefore, being Rows\_number\_Region equal to 21 it occurs that the participation of the regions is total and therefore several projects are located in the same region. Clearly, calculating the number of distinct projects participating in the relationship is equivalent to calculating the number of tuples present in the relationship itself and whose result was previously reported in table 23 (Rows\_number\_Project), since the CUPs are already the primary key of the Project table on which the foreign key Proj\_Region is integrated.

As can be seen from Figure 4, the projects have a partial participation in the 1:1 relationship with the beneficiary. Therefore, it is interesting to calculate the number of projects whose beneficiary is not known, by calculating the difference between the number of projects (Rows\_number\_Project) and the number of projects benefited, which we find through Query 3, and which is equal to 5224, i.e. the number of beneficiaries (Rows\_number\_Beneficiary). So, the number of non-beneficiary projects is equal to  $5246-5224=22$ .

#### Query 3.

```
SELECT COUNT(DISTINCT Proj_Beneficied)

FROM Beneficiary;
```

Similarly, the same analysis can be done with an Implementer, whereby being both "Rows\_Number\_Implementer" and the number of projects implemented equal to 5246, calculated with Query 4, we have further confirmation that a 1:1 relationship exists between Projects and Implementer, total on both sides.

#### Query 4.

```
SELECT COUNT(DISTINCT Project_Implemented)

FROM Implementer;
```

For what concern M:N relationships, such as the one between SDG\_Indicator and Sub-measure, it is possible to calculate the difference between all the Sub-measure and indicators in the database and those that actually participate in the relationship. Since the number of Indicators and that of Sub-measure is equal to 242 and 285 respectively, while the number of Indicators and Sub-Measure participating in the relation is equal to 35 and 199 respectively (calculated with Queries 5 and 6), it is easy to see the partial participation of both and it results that  $242-35=207$  indicators and  $285-199=86$  sub-measures do not participate in the relationship.

**Query 5.**

```
SELECT COUNT(DISTINCT COD_Submis_with_SDG)
FROM Submis_with_SDG;
```

**Query 6.**

```
SELECT COUNT(DISTINCT Indicator)
FROM Submis_with_SDG;
```

In the same way it is obtained that of the 285 Sub-measures only one does not participate in the M:N Relationship with the Goals (285-284) while all 1150 Goals participate in the relationship, where 284 and 1150 is the result that is found by executing the following Queries 7 and 8:

**Query 7.**

```
SELECT COUNT(DISTINCT COD_Submis_with_Goal)
FROM Submis_with_Goal;
```

**Query 8.**

```
SELECT COUNT(DISTINCT Goal_Submis)
FROM Submis_with_Goal;
```

Therefore, being the total number of rows of the report equal to 1489 (Rows\_number\_Submis\_with\_Goal) it is confirmed that the relationship is M:N with total participation in Goals and partial participation in Sub-measure.

As regards the 1:N relationship between Project and Sub-measure, the number of distinct Sub-measures present in the Project relationship is calculated using Query 9, which returns the value of 3 compared to the 5246 (non-distinct) Sub-measure present in the Project Table, calculated using Query 10.

**Query 9.**

```
SELECT COUNT(DISTINCT Proj_Submisura)
FROM Project;
```

**Query 10.**

```
SELECT COUNT(Proj_Submisura)
FROM Project;
```

Another way to query the database is to calculate the number of distinct values for the columns of each table if the variables are text type, while for numeric strings we have calculated the maximum and minimum to find the range of possible values. Queries 11, 12 and 13 show three examples of these cases.

**Query 11.**

```
SELECT COUNT(DISTINCT Grant_or_loand)
FROM Submisura;
```

**Query 42.**

```
SELECT MAX(Total_amount_submisura), MIN(Total_amount_submisura)
FROM Submisura;
```

**Query 53.**

```
SELECT MAX(Start_date), MIN(Start_date)
FROM Submisura;
```

Using the above queries, we repeated the same process for each column of each table and the results are shown in the Tables 24 and 25.

**Table 24.** Results of queries that analyze the distinct values

Table	Column	Distinct Values
<b>ADMINISTRATIONS</b>	Den_Adm	25
	Cod_Adm (PK)	25
<b>BENEFICIARY</b>	COD_Beneficiary (PK)	5224
	Des_Beneficiary	5215
	Proj_Beneficied	5224
<b>CITY</b>	City_Province	108
	COD_City (PK)	235
	Name_City	1760
	ZIP_CODE (PK)	1410

<b>COMPONENTE</b>	COD_Comp (PK)	16
	Des_Comp	16
	Missione_Comp	6
<b>GOAL</b>	COD_Goal (PK)	1150
	Name_Goal	1009
	Des_Goal	919
	ITA_or_UE	2
	Quarter	4
	Main_evidences	919
<b>IMPLEMENTER</b>	COD_Implementer (PK)	5246
	Des_Implementer	5237
	Project_Implemented	5246
<b>MILESTONE</b>	COD_Milestone (PK)	482
	Qual_Indicator	407
<b>MISSIONE</b>	COD_Missione (PK)	6
	Des_Missione	6
<b>MISURA</b>	COD_Misura (PK)	191
	Des_Misura	191
	Inv_or_Reform	2
	Comp_Misura	16
	Adm	24
<b>PROGRAMMER</b>	COD_Programmer (PK)	3
	Des_programmer	3
<b>PROJECT</b>	Cup (PK)	5246
	Name_Project	5246
	PNRR_Amount	994
	Proj_Programmer	3
	Proj_Submisura	3
	Proj_Region	21
<b>PROVINCE</b>	Province_Region	21
	COD_Province (PK)	108
	Name_Province	108
<b>REGION</b>	COD_Region (PK)	21
	Name_Region	21
<b>SDG_GOAL</b>	COD_SDG_Goal (PK)	17
	Des_SDG_Goal	17
<b>SDG_INDICATOR</b>	COD_SDG_Indicator (PK)	232
	Des_SGD_Indicator	232
	Target	169
<b>SDG_TARGET</b>	COD_SDG_Target (PK)	169
	Des_SDG_Target	168
	Goal	17
<b>SUBMIS_WITH_GOAL</b>	COD_Submis_with_Goal (PK)	284
	Submis_with_Goal	1150
<b>SUBMIS_WITH_SDG</b>	COD_Submis_with_SDG (PK)	199
	Indicator	35
<b>SUBMISURA</b>	COD_Submisura (PK)	285
	Des_Submisura	285
	Grant_or_loan	2
	Misura_Submisura	191
<b>TARGET</b>	COD_Target (PK)	668
	Measuring_unit	6

**Table 25.** Results of queries that analyze the MAX and MIN values

Table	Column	MAX	MIN
GOAL	Year	2027	2020
PROJECT	PNRR_Amount	114000000	100
SUBMISURA	Amount_proj	10255000000	0
	Amount_new_proj	6316960000	0
	Amount_FSC	2400000000	0
	Total_amount_submisura	13950000000	0
	End_date	31/12/2026	31/12/2022
	Start_date	01/01/2024	01/02/2020
TARGET	Starting_point	2130000000	0
	Final_Value	2769000000	-15

Eventually, another way to query the database is to ask how many nulls values there are in each column. Query 14 shows the example relating to the Total\_amount\_submisura column of the Submisura Table, while Table 26 shows the results relating to all tables.

```
SELECT COUNT(*) Total_amount_submisura
FROM Submisura
WHERE Total_amount_submisura IS NULL
```

**Table 26.** Results of queries that analyze the NULL values

Table	Column	NULL Values
ADMINISTRATIONS	COD_Adm (PK)	0
	Den_Adm	0
BENEFICIARY	COD_Beneficiary (PK)	0
	Des_Beneficiary	0
	Proj_Beneficiated	0
CITY	City_Province	0
	COD_City (PK)	0
	Name_City	0
	ZIP_CODE (PK)	0
COMPONENTE	COD_Comp (PK)	0
	Des_Comp	0
	Missione_Comp	0
GOAL	COD_Goal (PK)	0
	Name_Goal	0
	Des_Goal	0
	ITA_or_UE	0
	Quarter	0
	Year	0
	Main_evidences	1054
IMPLEMENTER	COD_Implementer (PK)	0
	Des_Implementer	0

	Project_Implemented	0
<b>MILESTONE</b>	COD_Milestone (PK)	0
	Qual_Indicator	4
<b>MISSIONE</b>	COD_Missione (PK)	0
	Des_Missione	0
<b>MISURA</b>	COD_Misura (PK)	0
	Des_Misura	0
	Inv_or_Reform	0
	Comp_Misura	0
<b>PROGRAMMER</b>	COD_Programmer (PK)	0
	Des_programmer	0
<b>PROJECT</b>	Cup (PK)	0
	Name_Project	0
	PNRR_Amount	0
	Proj_Programmer	0
	Proj_Submisura	0
	Proj_Region	0
<b>PROVINCE</b>	Province_Region (PK)	0
	COD_Province	0
	Name_Province	0
<b>REGION</b>	COD_Region (PK)	0
	Name_Region	0
<b>SDG_GOAL</b>	COD_SDG_Goal (PK)	0
	Des_SDG_Goal	0
<b>SDG_INDICATOR</b>	COD_SDG_Indicator (PK)	0
	Des_SGD_Indicator	0
	Target	0
<b>SDG_TARGET</b>	COD_SDG_Target (PK)	0
	Des_SDG_Target	0
	Goal	0
<b>SUBMIS_WITH_GOAL</b>	COD_Submis_with_Goal (PK)	0
	Submis_with_Goal	0
<b>SUBMIS_WITH_SDG</b>	COD_Submis_with_SDG (PK)	0
	Indicator	0
<b>SUBMISURA</b>	COD_Submisura (PK)	0
	Des_Submisura	59
	Grant_or_loan	0
	Amount_proj	0
	Amount_new_proj	0
	Amount_FSC	0
	Total_amount_submisura	0
	Start_date	86
	End_date	86
	Misura_Submisura	0
<b>TARGET</b>	COD_Target (PK)	0
	Measuring_unit	0
	Starting_point	0
	Final_Value	0



## Additional queries

### 1. Count the Total Amount for each Mission and order it from the highest to the smallest.

```
SELECT SUM(Total_Amount_submisura ) AS 'Total_Amount_Missione', COD_Missione
FROM Submisura, Misura, Componente, Missione
WHERE Misura_Submisura=COD_Misura AND Comp_Misura=COD_Comp AND
Missione_Comp=COD_Missione
GROUP BY COD_Missione
ORDER BY Total_Amount_Missione ASC;
```

Total_Amount_Missione	COD_Missione
15625541084	M6
19850900000	M5
25396732501	M3
30876000000	M4
40291453254	M1
59458551051	M2

## 2. Count the number of the distinct Submission for each SDG Goal

```
SELECT COUNT(DISTINCT COD_Submis_with_SDG) AS 'Number_of_Submis_with_SDG_Goal',  
COD_SDG_Goal  
FROM Submis_with_SDG, SDG_Indicator, SDG_Target, SDG_Goal  
WHERE Indicator=COD_SDG_Indicator AND Target=COD_SDG_Target AND Goal=COD_SDG_Goal  
GROUP BY COD_SDG_Goal  
ORDER BY Number_of_Submis_with_SDG_Goal DESC;
```

Number_of_Submis_with_SDG_Goal	COD_SDG_Goal
71	9
40	13
35	7
23	8
21	4
20	11
12	1
12	3
9	12
6	6
4	5
3	16
1	14

## 3. Count the Total Amount of the Projects for each Region

```
SELECT Name_Region, SUM(PNRR_Amount) AS 'Total_Amount'  
FROM Project, Region  
WHERE Proj_Region=COD_Region  
GROUP BY COD_Region  
ORDER BY Name_Region ASC;
```

Name_Region	Total_Amount
ABRUZZO	38123240
AMBITO NAZIONALE	114000000
BASILICATA	3414910
CALABRIA	20146170
CAMPANIA	110443230
EMILIA-ROMAGNA	84858500
FRIULI-VENEZIA GIULIA	21361900
LAZIO	88201490
LIGURIA	19451860
LOMBARDIA	205856660
MARCHE	39246020
MOLISE	9465670
PIEMONTE	76544980
PUGLIA	63112700
SARDEGNA	16634480
SICILIA	52985907
TOSCANA	88624830
TRENTINO-ALTO ADIGE/SUDTIROL	17503120
UMBRIA	6427110
VALLE D'AOSTA	5186000
VENETO	105536280

#### 4. Count the Total Amount of the projects for Lecce, Brindisi and Taranto

```

SELECT Name_Province, SUM(PNRR_Amount) AS 'Total_Amount'
FROM Project, Province
WHERE Proj_Province=COD_Province AND Name_Province IN ('LECCE', 'BRINDISI', 'TARANTO')
GROUP BY Name_Province;

```

Name_Province	Total_Amount
TARANTO	1621310
BRINDISI	1707960
LECCE	3679050

## 5. Count the number of the Submeasures already started

```
SELECT COUNT(DISTINCT COD_Submisura) AS 'Number_of_submis_already_started'
FROM Submisura
WHERE Start_date BETWEEN '2020-02-21' AND '2022-12-08';
```

Number_of_submis_already_started
177

## 6. Counts the number of projects managed by each Administration that manages at least one project

```
SELECT Den_Adm, COUNT(DISTINCT CUP) AS 'Projects_managed'
FROM Project, Administrations, Submisura, Misura
WHERE Proj_Submisura=COD_Submisura AND Misura_Submisura=COD_Misura AND Adm=COD_Adm
GROUP BY Den_Adm
ORDER BY Den_Adm ASC;
```

Den_Adm	Projects_managed
MINISTERO DEL TURISMO	1
MINISTERO DELLO SVILUPPO ECONOMICO	5224
Ministro della Pubblica Amministrazione	21

**7. Count the total amount, the highest amount, the lowest amount and the average of the Project PNRR Amount**

```
SELECT SUM(PNRR_Amount) AS Total_Amount, MAX(PNRR_Amount) AS Highest_Amount,  
MIN(PNRR_Amount) AS Lowest_Amount, AVG(PNRR_Amount) AS Average_Amount  
FROM Project;
```

Total_Amount	Highest_Amount	Lowest_Amount	Average_Amount
1187125057	114000000	100	226291,471

**8. Count the number of the projects that have the amount higher than the Average Amount**

```
SELECT COUNT(DISTINCT CUP) AS 'Project_with_amount_higher_than_the_Avarage_Amount'  
FROM Project  
WHERE PNRR_Amount > ALL (SELECT AVG(PNRR_Amount)  
FROM Project);
```

Project_with_amount_higher_than_the_Avarage_Amount
1482

**9. Show the Component with the highest number of projects, its description and its number of projects**

```
SELECT COD_Comp, Des_Comp, COUNT(DISTINCT CUP) AS 'Number_of_proj'  
FROM Project, Submisura, Misura, Componente  
WHERE Proj_Submisura=COD_Submisura AND Misura_Submisura=COD_Misura AND  
Comp_Misura=COD_Comp  
GROUP BY COD_Comp  
ORDER BY Number_of_proj DESC
```

LIMIT 1;

COD_Comp	Des_Comp	Number_of_proj
M1C2	Digitalizzazione, innovazione e competitività nel sistema produttivo	5224

**10. Retrieve the name, the CUP and the total amount of the projects located in Lecce**

SELECT Name\_Project, CUP, PNRR\_Amount AS 'Total\_amount'

FROM Project, City

WHERE Proj\_City=COD\_City AND Name\_City='LECCE'

Name_Project	CUP	Total_amount
TECNICA SRL*FINANZIAMENTI PER L'INTERNAZIONALIZZAZIONE DELLE IMPRESE E PER LA TRANSIZIONE DIGITALE*VIA DEGLI INTARSIATORI ROLESI 1	E11B21008800006	300000
PASTIFICIO GENTILE SRL*FINANZIAMENTI PER L'INTERNAZIONALIZZAZIONE DELLE IMPRESE E PER LA TRANSIZIONE DIGITALE*VIA CASTELLO 12	E11B21010940006	300000
CUOIERIA FIORENTINA SRL*FINANZIAMENTI PER L'INTERNAZIONALIZZAZIONE DELLE IMPRESE E PER LA TRANSIZIONE DIGITALE*VIA DEI CILIEGI 25	E11B21012680006	300000

## REFERENCES

(2021) The National Recovery and Resilience Plan (PNRR). In: MEF.

<https://www.mef.gov.it/en/focus/The-National-Recovery-and-Resilience-Plan-NRRP/>

Elmasri R, Navathe S (2015) Fundamentals of Database Systems. Pearson, Hoboken, NJ

[www.governo.it](http://www.governo.it)

[www.italiandomani.gov.it](http://www.italiandomani.gov.it)

[www.villaggioinformatico.it](http://www.villaggioinformatico.it)