

# Data Visualization From a Category Theory Perspective

Davi Sales Barreira, Asla Medeiros e Sá

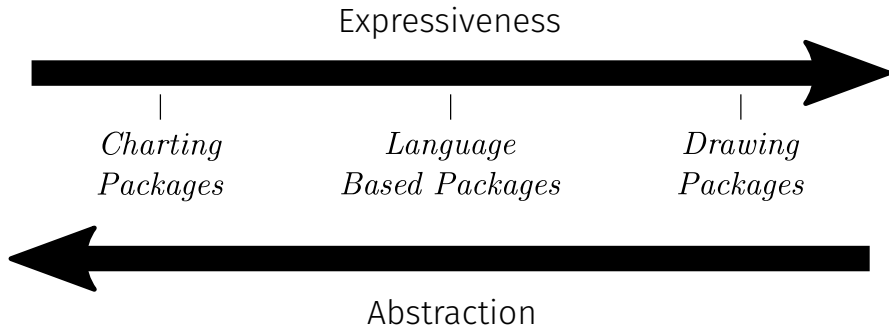
FGV - EMAp, IMPATech

February 5, 2025

# Motivation

---

Balance expressiveness and abstraction in data visualization frameworks.



# Motivation

How can we represent complex visualizations without resorting to low-level specifications?

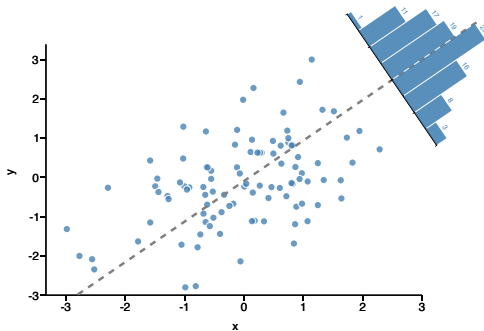


Figure: Rotated histogram aligned with second main PCA axis.

# Motivation

How can we represent complex visualizations without resorting to low-level specifications?

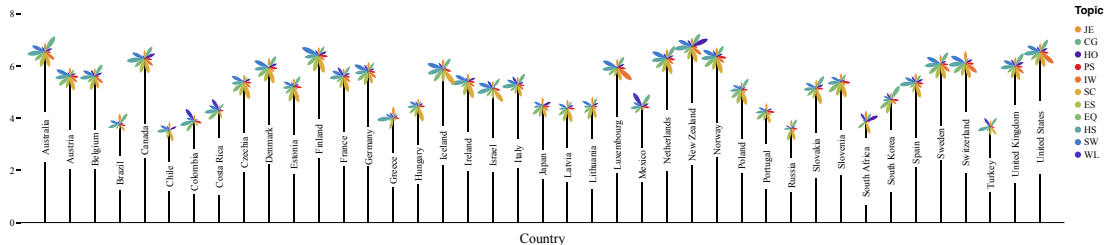


Figure: OECD Better Life Index visualization Stefaner and OECD [3].

# Overview

---

- **Day 1:** Basics of Category Theory
- **Day 2:** Programming with Category Theory
- **Day 3:** Data Visualization Theory
- **Day 4:** Data Visualization + Categorical Programming

# Table of contents

---

1. What are Categories?
2. Examples of Categories
3. Universal Constructions
4. Functors
5. Natural Transformations
6. Monoids and Monads
7. F-Algebras and F-Coalgebras

# Why Category Theory?

---

**Category Theory** is a branch of mathematics that studies general abstract structures through their relationships.

**Origin:** Samuel Eilenberg e Saunders Mac Lane - 1940

As pointed by Fong and Spivak [2], Category Theory is unmatched in its ability to organize and relate abstractions.

## Category Theory

**Mathematics**

**Programming**

**Data Visualization**

# What are Categories?

---

Category Theory is a branch of mathematics that studies general abstract structures through their relationships.

## Definition (Category)

- A collection of objects  $\text{Ob}_{\mathcal{C}}$ .
- A collection of morphisms  $\text{Hom}_{\mathcal{C}}$ , where each morphism has a source object and a target object.  $\text{Hom}_{\mathcal{C}}(A, B)$  is the collection of morphisms going from object  $A$  to object  $B$ .
- A binary operation  $\circ : \text{Hom}_{\mathcal{C}}(A, B) \times \text{Hom}_{\mathcal{C}}(B, C) \rightarrow \text{Hom}_{\mathcal{C}}(A, C)$  such that:
  1. **Associative:**  $(h \circ g) \circ f = h \circ (g \circ f)$ .
  2. **Identity:** Every object has an identity morphism  $1_A \in \text{Hom}_{\mathcal{C}}(A, A)$ .



# What are Categories?

---

## Definition (Small and Locally Small Category)

A category  $\mathcal{C}$  is *small* if  $\text{Ob}_{\mathcal{C}}$  and  $\text{Hom}_{\mathcal{C}}$  are sets. A category  $\mathcal{C}$  is *locally small* if for any  $A, B \in \text{Ob}_{\mathcal{C}}$ , then  $\text{Hom}_{\mathcal{C}}(A, B)$  is a set. Note that a small category is also locally small.

Note that when talking about  $\text{Ob}_{\mathcal{C}}$  and  $\text{Hom}_{\mathcal{C}}$ , we didn't say that they were sets, instead we called them *classes*. The reason for this lies in the foundations of Set Theory. There are collections in mathematics that are “larger” than sets, e.g. the “set” of all sets, which itself cannot be a set, otherwise it would incur in a paradox (Russell's Paradox). A way to deal with this is making a distinction between classes and sets. This point is quite technical; readers interested in understanding this nuance can check books such as Borceux [1].

# Examples of Categories

---

The category **1** consists of  $\text{Ob}_{\mathbf{1}} := \{A\}$  and  $\text{Hom}_{\mathbf{1}} = \text{id}_A$ .



## Examples of Categories

---

The category **2** consists of  $\text{Ob}_2 := \{A, B\}$  and  $\text{Hom}_1 = \{id_A, id_B, f\}$ , where  $f : A \rightarrow B$ . The diagram for such category is shown below.

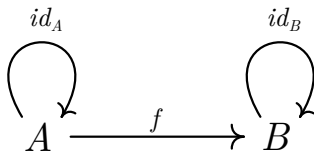


Figure: Category 2.

## Examples of Categories

The category **3** has three morphisms besides the identities. The morphisms are  $f$ ,  $g$  and their composition  $g \circ f$ . The figure below illustrates the category with all its morphisms.

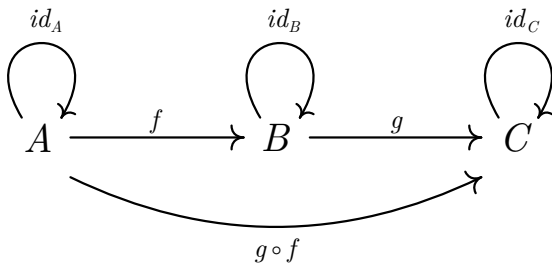


Figure: Category **3** showing all morphisms.

# Examples of Categories

---

When drawing categories, it is common to omit the identity and/or composition morphism. From here on, we do the same, whenever the context is clear.

$$A \xrightarrow{f} B \xrightarrow{g} C$$

Figure: Category **3** omitting morphisms.

# Examples of Categories

---

The discrete category  $\underline{\mathbf{N}}$  is the category with  $N$  objects and  $\text{Hom}_{\underline{\mathbf{N}}} := \{id_1, \dots, id_N\}$ . An example of this category is illustrated below.



Figure: Category  $\underline{\mathbf{N}}$ .

# Examples of Categories

Given a category  $\mathcal{C}$  and an object  $S$  of this category, we can define a slice category  $\mathcal{C}/S$ , where:

- The objects are tuples  $(A, f)$  where  $A$  is an object in  $\mathcal{C}$  and  $f : A \rightarrow S$  is a morphism.
- A morphism  $\varphi_{(A,B)} : (A, f) \rightarrow (B, g)$  is equivalent to a morphism  $\varphi \in \text{Hom}_{\mathcal{C}}(A, B)$  such that  $f = g \circ \varphi$ .

$$\begin{array}{ccc} & \mathcal{C}/S & \\ (A, f) & \xrightarrow{\phi_{A,B}} & (B, g) \end{array} \cong \begin{array}{ccc} & \mathcal{C} & \\ A & \xrightarrow{\phi} & B \\ & \searrow f \quad \swarrow g & \\ & S & \end{array}$$

Figure: Example of slice category.

# Examples of Categories

---

## Definition (Subcategory)

Let  $\mathcal{C}$  be a category. A *subcategory*  $\mathcal{S}$  of  $\mathcal{C}$  is such that

- (i)  $\text{Ob}_{\mathcal{S}} \subseteq \text{Ob}_{\mathcal{C}}$ ;
- (ii) For every  $A, B \in \text{Ob}_{\mathcal{S}}$ , we have  $\text{Hom}_{\mathcal{S}}(A, B) \subseteq \text{Ob}_{\mathcal{C}}(A, B)$ ;
- (iii) Composition and identity in  $\mathcal{S}$  are the same as in  $\mathcal{C}$ , restricted to morphisms and objects of  $\mathcal{S}$ .

A subcategory  $\mathcal{S}$  is said to be *wide* if  $\text{Ob}_{\mathcal{S}} = \text{Ob}_{\mathcal{C}}$ , and it is said to be *full* if for every  $A, B \in \text{Ob}_{\mathcal{S}}$ , then  $\text{Hom}_{\mathcal{S}}(A, B) = \text{Ob}_{\mathcal{C}}(A, B)$ . Finally, a subcategory is *thin* if for every  $A, B \in \text{Ob}_{\mathcal{S}}$  the set  $\text{Hom}_{\mathcal{S}}(A, B)$  has only a single morphism.



# Examples of Categories

## Definition (Dual Category)

Given a category  $\mathcal{C}$ , the dual (opposite) category  $\mathcal{C}^{op}$  is defined as:

- $\text{Ob}_{\mathcal{C}^{op}} = \text{Op}_{\mathcal{C}}$ ;
- For every morphism  $f : A \rightarrow B$  in  $\mathcal{C}$ , have an equivalent morphism  $f^{op} : B \rightarrow A$  in  $\mathcal{C}^{op}$ ;
- The composition in  $\mathcal{C}^{op}$  satisfies  $f^{op} \circ g^{op} = (g \circ f)^{op}$ .

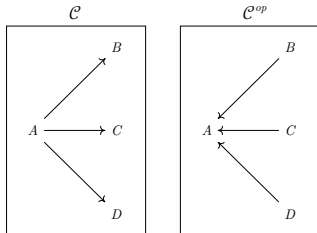


Figure: Example of dual category.

# Examples of Categories

---

Here are some more interesting categories:

1. **Set** which is the category of sets, where the objects are sets and the morphisms are functions between sets.
2. **Top** is the category where topological spaces are the objects and continuous functions are the morphisms.
3. **Vec $_{\mathbb{F}}$**  is the category where vector spaces over field  $\mathbb{F}$  are the objects, and linear transformations are the morphisms.
4. **Mon** is the category of monoids, where morphisms are monoid homomorphisms.

# Isomorphisms, monomorphism and epimorphism

A very important definition in Category Theory is the notion of isomorphism. In Set Theory, we say that two sets are isomorphic if there is a bijective function between them. Yet, this concept is not restricted to Set Theory and can be generalized in Category Theory as follows:

## Definition (Categorical Isomorphism)

Let  $\mathcal{C}$  be a category with  $X, Y \in \text{Ob}_{\mathcal{C}}$  and  $f \in \text{Hom}_{\mathcal{C}}(X, Y)$ .

- (i) We say that  $f$  is *left invertible* if there exists  $f_l \in \text{Hom}_{\mathcal{C}}(Y, X)$  such that  $f_l \circ f = \text{id}_X$ ;
- (ii) We say that  $f$  is *right invertible* if there exists  $f_r \in \text{Hom}_{\mathcal{C}}(Y, X)$  such that  $f \circ f_r = \text{id}_Y$ ;
- (iii) We say that  $f$  is invertible if it's both left and right invertible.

**Exercise:** Prove that in the category **Set**, the categorical isomorphism is equivalent to the set theoretic isomorphism.

# Isomorphisms, monomorphism and epimorphism

## Definition (Monomorphism)

In a given category  $\mathcal{C}$ , a morphism  $m : X \rightarrow Y$  is said to be a monomorphism (monic), if for every  $A \in \text{Ob}_{\mathcal{C}}$  and  $f, g : A \rightarrow X$  we have that:

$$m \circ f = m \circ g \implies f = g.$$

## Definition (Epimorphism)

In a given category  $\mathcal{C}$ , a morphism  $e : X \rightarrow Y$  is said to be an epimorphism (epic), if for every  $A \in \text{Ob}_{\mathcal{C}}$  and  $f, g : A \rightarrow X$  we have that:

$$f \circ e = g \circ e \implies f = g.$$

# Isomorphisms, monomorphism and epimorphism

## Proposition

The following properties on monomorphism and epimorphism are true:

1.  $f$  left-invertible  $\implies f$  is monic. The converse is not true.
2.  $f$  right-invertible  $\implies f$  is epic. The converse is not true.
3.  $f$  invertible  $\implies f$  is monic and epic. The converse is not true.
4.  $f$  monic and right-invertible  $\implies f$  is isomorphism.
5.  $f$  epic and left-invertible  $\implies f$  is isomorphism.

## Proof.

Left as exercise.



# Universal Constructions

---

Objects defined in terms of existence and uniqueness of morphisms are known as **universal constructions**.

## Definition (Zero, Initial and Terminal)

Let  $\mathcal{C}$  be a category.

1. An object  $I \in \text{Ob}_{\mathcal{C}}$  is *initial* if for every  $A \in \text{Ob}_{\mathcal{C}}$ , there is exactly one morphism from  $I$  to  $A$ . Thus, from  $I$  to  $I$  there is only the identity.
2. An object  $T \in \text{Ob}_{\mathcal{C}}$  is *terminal* if for every  $A \in \text{Ob}_{\mathcal{C}}$ , there is exactly one morphism from  $A$  to  $T$ . Thus, from  $I$  to  $I$  there is only the identity.
3. An object is *zero* if it is both terminal and initial.

# Universal Constructions

---

## Theorem

*Every initial object is unique up to an isomorphism, i.e. if in a category there are two initial objects, then they are isomorphic. Similarly, terminal objects are unique up to an isomorphism. Moreover, the isomorphism is unique between initial object, and between terminal objects.*

## Proof.

Left as exercise.



# Universal Constructions

---

In Set Theory, we are used to the notion of a Cartesian product. Similarly to how we did for isomorphism, the idea of a product can be generalized via Category Theory

## Definition (Product)



# Set Category

---

A very important category is **Set**. Here are some of its properties:

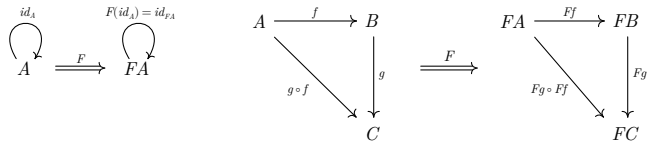
- Initial object:  $\emptyset$ ;
- Terminal object: any singleton set up to isomorphism;
- For any two objects  $A$  and  $B$ , the  $\text{Hom}(A, B)$  is also an object (called exponential object  $B^A$ );
- Monomorphisms are equivalent to injective functions;
- Epimorphisms are equivalent to surjective functions.

# Category Theory Brief Introduction

## Definition (Functor)

Let  $\mathcal{C}$  and  $\mathcal{D}$  be two categories. A functor  $F : \mathcal{C} \rightarrow \mathcal{D}$  is a pair of mappings with the following properties:

Covariant Functor



Contravariant Functor

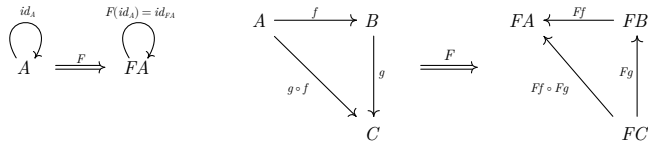


Figure: Diagrams showcasing the properties of functors.

# Category Theory Brief Introduction

## Definition (Natural Transformations)

Let  $\mathcal{C}$  and  $\mathcal{D}$  be categories, and let  $F, G : \mathcal{C} \rightarrow \mathcal{D}$  be functors. A natural transformation  $\alpha : F \rightarrow G$  is such that the following diagram commutes:

$$\begin{array}{ccc} FA & \xrightarrow{\alpha_A} & GA \\ \downarrow Ff & & \downarrow Gf \\ FB & \xrightarrow{\alpha_B} & GB \end{array}$$

$$Gf \circ \alpha_A = \alpha_B \circ Ff$$

**Figure:** Commutative diagram of a natural transformation highlighting the commutative property of the definition.

# Category Theory Brief Introduction

---

**Monoids** and **Monads** are two ubiquitous constructions both in Category Theory and Functional Programming. These two concepts will be used when talking about data visualization. Therefore, it is required of us to introduce these constructions.

Let's start with the definition of a monoid in the context of Set Theory.

## Definition (Monoid - Set Theory)

A monoid is a triple  $(M, \otimes, e_M)$  where  $M$  is a set,  $\otimes : M \times M \rightarrow M$  is a binary operation and  $e_M$  the neutral element, such that:

1.  $a \otimes (b \otimes c) = (a \otimes b) \otimes c$
2.  $a \otimes e_M = e_M \otimes a = a$ .

An example of a monoid is  $(\mathbb{N} \cup \{0\}, +, 0)$ . It is easy to check that the summation operator satisfies the associativity neutrality properties.

# Category Theory Brief Introduction

## Definition (Monoid in the category **Set**)

A monoid in **Set** is a triple  $(M, \mu, \eta)$ , where  $M \in \text{Ob}_{\mathbf{Set}}$ ,  $\mu : M \times M \rightarrow M$  and  $\eta : 1 \rightarrow M$  are two morphisms in **Set** satisfying the commutative diagrams below. Note that  $1$  is the terminal object in **Set**, i.e. the singleton set (which is unique up to an isomorphism).

The figure contains two commutative diagrams. The left diagram is a square with vertices  $M \times M \times M$  (top-left),  $M \times M$  (top-right),  $M \times M$  (bottom-left), and  $M$  (bottom-right). The top horizontal arrow is  $id_M \times \mu$ , the bottom horizontal arrow is  $\mu$ , the left vertical arrow is  $\mu \times id_M$ , and the right vertical arrow is  $\mu$ . The right diagram is a triangle with vertices  $M$  (top-left),  $M \times M$  (top-right), and  $M$  (bottom). The top-left horizontal arrow is  $\eta \times id_M$ , the top-right horizontal arrow is  $id_M \times \eta$ , the left diagonal arrow is  $id_M$ , the right diagonal arrow is  $id_M$ , and the vertical arrow from  $M \times M$  to  $M$  is  $\mu$ .

Figure: Commutative diagram for monoid.

# Category Theory Brief Introduction

## Definition (Monad)

A monad is a monoid in  $\mathbf{End}_{\mathcal{C}}$ , which is the triple  $(T, \mu, \eta)$ , where  $T : \mathcal{C} \rightarrow \mathcal{C}$  is a functor,  $\mu : T \circ T \rightarrow T$  and  $\eta : 1 \rightarrow T$  are natural transformations in  $\mathbf{End}_{\mathcal{C}}$  satisfying the commutative diagrams below. Note that  $1$  is the identity functor in  $\mathcal{C}$ .

$$\begin{array}{ccc} T \circ T \circ T & \xrightarrow{id_T * \mu} & T \circ T \\ \downarrow \mu * id_T & & \downarrow \mu \\ T \circ T & \xrightarrow{\mu} & T \end{array} \qquad \begin{array}{ccccc} T & \xrightarrow{\eta * id_T} & T \circ T & \xleftarrow{id_T * \eta} & T \\ & \searrow id_T & \downarrow \mu & \swarrow id_T & \\ & & T & & \end{array}$$

Figure: Commutative diagram for monad.

# References

---

- [1] Borceux, F. (1994). *Handbook of categorical algebra: volume 1, Basic category theory*, volume 1. Cambridge University Press.
- [2] Fong, B. and Spivak, D. I. (2019). *An invitation to applied category theory: seven sketches in compositionality*. Cambridge University Press.
- [3] Stefaner, M. and OECD (2012). Oecd better life index. <http://www.oecdbetterlifeindex.org/>. Accessed on 14 Oct. 2012.

# The End