

$$(X, O) = e^{-\frac{x^2}{2\sigma^2}}$$

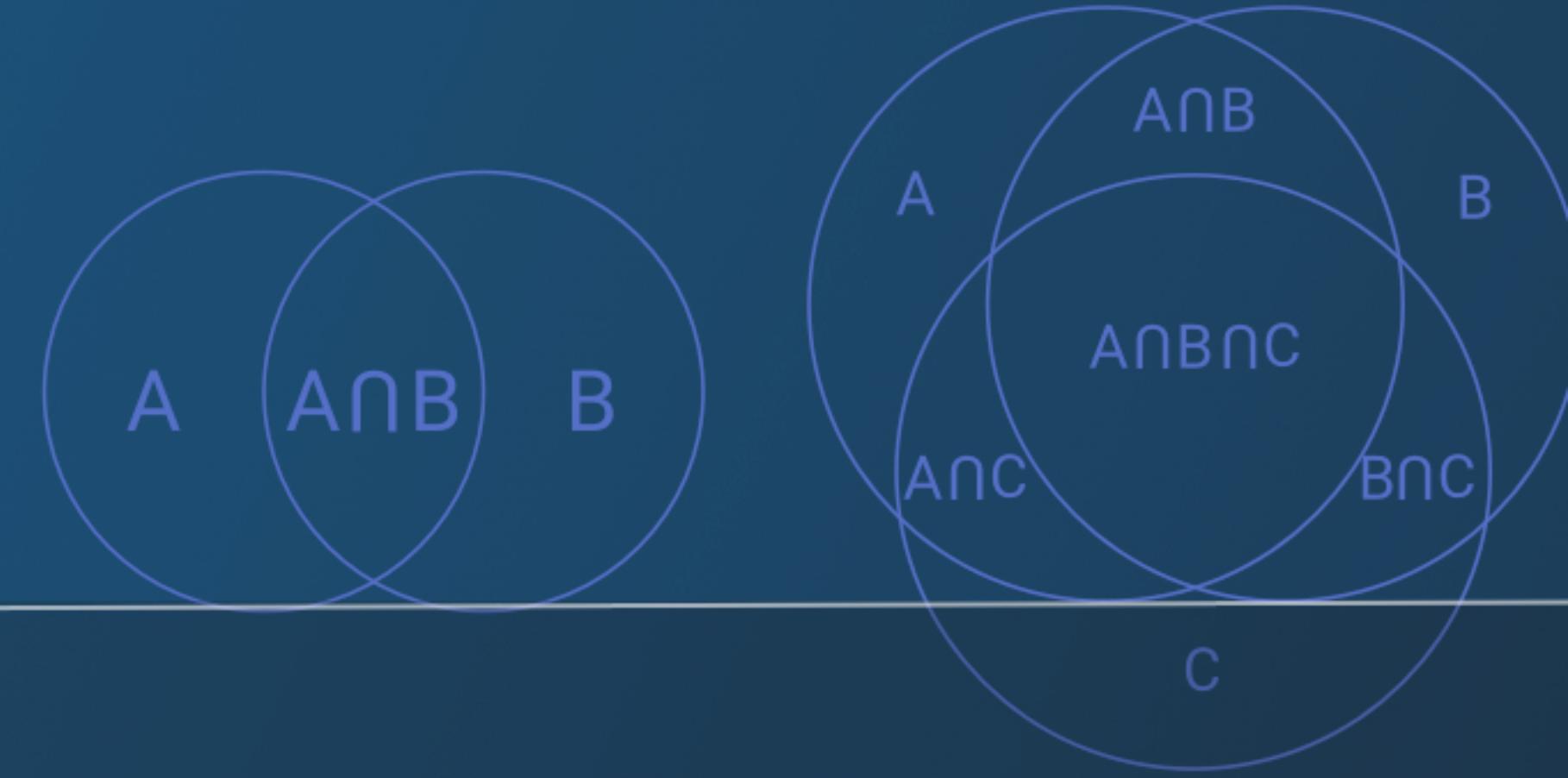
$$x(X, O) = -\frac{x}{\sigma^2} G(X, O) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

$$xx(X, O) = \frac{x^2 - \sigma^2}{\sigma^4} G(X, O) = \frac{x^2 - \sigma^2}{\sigma^4} e^{-\frac{x^2}{2\sigma^2}}$$

$$xxx(X, O) = -\frac{x^3 - x\sigma^2}{\sigma^6} G(X, O) = -\frac{x^3 - x\sigma^2}{\sigma^6} e^{-\frac{x^2}{2\sigma^2}}$$

# Julia 程式語言學習馬拉松

## Day 20



$$\ln(x + \sqrt{1+x^2}) + x - \frac{1}{x + \sqrt{1+x^2}} \left( 1 + \frac{x}{\sqrt{1+x^2}} \right)$$



cupay 陪跑專家 : James Huang

$$\begin{aligned} &= \frac{1}{x + \sqrt{1+x^2}} \left( \frac{\sqrt{1+x^2} + x}{\sqrt{1+x^2}} \right) + \frac{1}{\sqrt{1+x^2}} - \frac{x^2}{\sqrt{(1+x^2)^3}} \\ &= \frac{1}{\sqrt{1+x^2}} + \frac{1}{\sqrt{1+x^2}} - \frac{x^2}{\sqrt{(1+x^2)^3}} \\ &= \frac{2}{\sqrt{1+x^2}} - \frac{x^2}{\sqrt{(1+x^2)^3}} = \frac{2(1+x^2) - x^2}{\sqrt{(1+x^2)^3}} = \frac{2+x^2}{\sqrt{(1+x^2)^3}} \end{aligned}$$

# DataFrames.jl 介紹 (四): Categorical Data 與 Missing Data





# 重要知識點



- 在現實世界中，我們獲取的資料內常包含類別資料，也常可能在資料中會有缺值的情況，DataFrames.jl 及 Julia 本身即提供了對於類別資料與缺值處理的功能。
- 搭配 CategoricalArrays.jl 套件，更方便地提供對於類別資料與缺值的操作與處理。



# CategoricalArrays.jl



- 建立 CategoricalArray 的方式有兩種：
  - 透過建構子 CategoricalArray()
  - 透過 categorical() 函式
- 呼叫 levels() 產生 CategoricalArray 中可出現的類別；與 Base.unique() 不同的是，unique 會包含 missing，但是 levels() 不會。
- compress() 函式可以壓縮 CategoricalArray 以節省記憶體的用量.



# 將字符串 (String) 欄位轉換為 CategoricalString



- 在 DataFrame 中如果需要將定欄位轉換為類別型態的話，可以使用 categorical!() 函式。如下例將 industry 欄位從 String 轉換為 CategoricalString。

```
1 categorical!(df, :industry, compress=true)
```

5 rows × 3 columns

	index	name	industry
	Int64	String	Categorical...
1	1	James	製造業
2	2	Jim	營建工程業
3	3	Adam	批發及零售業
4	4	Dawn	資通訊服務業
5	5	Jason	金融及保險業



# Missing 的處理



- Julia 提供 Missing 型別來代表當資料中有缺值的情況發生時，類比於 SQL 的 NULL、R 語言中的 NA。
- 使用 Base.skipmissing() 函式，可將 vector 中的 missing 值去掉。
- 呼叫 collect() 函式可以取出非 missing 值。
- 配合 skipmissing()，呼叫 Base.coalesce() 以其他值來取代 missing。
- 在 DataFrames.jl 中提供 dropmissing() 與 dropmissing!(), 可將 DataFrame 中含有 missing 的 row 去掉。兩者不同點在於，dropmissing!() 會 in-place 改變 DataFrame.

# 知識點 回顧

- 今天的內容介紹了 DataFrame 搭配 CategoricalArrays 套件，可以運用在資料科學的任務，例如資料清洗、特徵工程中。
- 在現實世界中，我們獲取的資料內常包含類別資料，也常可能在資料中會有缺值的情況，均可使用這些套件來協助我們處理。



## 推薦閱讀

- [CategoricalArrays.jl 官方文件](#)





解題時間

請跳出 PDF 至官網 Sample Code  
& 作業開始解題