

# Generar nombres de empresas biotecnológicas al azar

Agustín Castro

2023-10-13

## Contents

Píldoras_R. Material de formación . . . . .	1
Vectores con nombres de empresas y adjetivos. . . . .	1
Comprobar número de elementos en los vectores . . . . .	2
Buscar elementos únicos y duplicados . . . . .	3
Función para generar $n$ nombres . . . . .	4

## Píldoras\_R. Material de formación

En esta práctica se generan nombres de empresas biotecnológicas al azar, a partir de dos listados de elementos obtenidos de consultas a chatgpt.

[mi\_blog] <https://agustincastro.es>

## Vectores con nombres de empresas y adjetivos.

Los nombres de las empresas estarán formados por dos palabras, un nombre, almacenado en **lab\_names**, y un adjetivo, en **lab\_adjectives**. Ambos vectores han sido obtenidos de dos consultas a **chatgpt**. Se le pidió un listado con 150 posibles nombres de empresas biotecnológicas que estuvieran formados por una única palabra. Igualmente, se hizo lo mismo para el vector con los “adjetivos”.

```
lab_names <- c("BioSynth", "ChemixLab", "PharmaWave", "LifeCore", "NanoBio", "GenoTechnics",
               "QuantumChem", "BioGene", "MediSynth", "LabPro", "Bioteched", "EnviroLab",
               "SynthCorp", "BioPharma", "GenoTech", "ChemWare", "MediGene", "QuantumBio",
               "LifeLab", "NanoTech", "SynthLife", "ChemixGen", "BioInno", "PharmaSyn",
               "GenoLab", "MediLife", "BiotechLab", "LabBio", "EnviroGen", "NanoPharm",
               "ChemCore", "Lifer", "QuantumCore", "SynthTech", "MediPharm", "GenoWave",
               "BioGenix", "LabSynth", "PharmaCorp", "NanoCore", "ChemBio", "EnviroTech",
               "LifeTech", "QuantumGen", "MediGenix", "BioCorp", "GenoBio", "SynthWave",
               "BiotechX", "LabGen", "NanoLife", "PharmaInno", "EnviroBio", "ChemPro",
               "LifeInno", "QuantumPharm", "MediPro", "GenoInno", "BioTech", "SynthPharm",
               "LabPharm", "NanoSynth", "LifeBio", "EnviroSynth", "MediSynthX", "GenoPro",
               "ChemSynth", "BioPharmX", "QuantumLab", "NanoPharma", "LifePharm", "SynthInno",
               "MediPharma", "GenoCorp", "BioWare", "LabLife", "EnviroInno", "NanoGenix",
               "ChemProX", "QuantumBios", "MediCore", "SynthLifeX", "BioLifeX", "PharmaTech",
               "LifeGenix", "GenoPharmaX", "NanoBioX", "ChemTechX", "MediBioX", "SynthPharmaX",
               "LabGenix", "EnviroGenix", "NanoGen", "BioGen", "LifeProX", "MediPharmX", "GenoGenixX",
```

```

"BioSolve", "ChemBioX", "GenoNova", "NanoCoreX", "SynthPharmaZ", "LabInnoX",
"EnviroLifeZ", "MediTechX", "QuantumProZ", "LifeGenius", "PharmaBioZ",
"BioTechLabZ", "GenoScienceZ", "NanoPharmX", "ChemXcelZ", "SynthoMedX",
"MediWaveZ", "LabPulseX", "EnviroTechZ", "GenoSynthX", "BioMaxZ",
"NanoBioTechZ", "LifeTechX", "QuantumBioX", "ChemWaveZ", "SynthoLifeX",
"LabCoreZ", "GenoTechX", "EnviroPharmX", "BioInnoZ", "NanoProX", "LifePharmZ",
"PharmaGenX", "QuantumCoreX", "MediSynthZ", "ChemXpertX", "SynthoGenX",
"LabXpressZ", "GenoInnoZ", "EnviroBioX", "BioTechXcelX", "NanoPharmaZ",
"LifeGenixX", "PharmaTechZ", "QuantumBioXcelX", "MediWaveXcelX", "ChemCoreXcelZ",
"SynthoPharmaX", "LabTechZ", "BioGenXcel", "NanoCoreInno", "LifePharmTech",
"SynthoMedInnovations", "QuantumGen")

lab_adjetives <- c("Innovative", "Technological", "Biologicals", "Therapeutic", "Pharmacological",
"Biogenetic", "Clinical", "Modern", "Efficient", "Investigative",
"Experimental", "Biomedical", "Molecular", "Engineering",
"Precise", "Analyticals", "Synthetics", "Futuristic", "Personalized",
"Specialized", "Genetic", "Nano", "Neuropharmaceutical",
"Ecological", "Veterinary", "International", "Scientific",
"Inspiring", "Sustainable", "Advanced", "Revolutionary", "Healthy",
"Radiological", "Biochemicals", "Cloning", "Environmentaly", "Regenerative",
"Bioinformatics", "Microscopic", "Pioneering", "Metabolic",
"Pharmaceuticals", "Medical", "Intelligent", "Genomic", "Chemical", "Analytical",
"Organic", "Inorganic", "Biochemical", "Physical", "Nuclear", "Synthetic",
"Molecular", "Polymer", "Atomic", "Pharmaceutical", "Environmental",
"Quantum", "Spectroscopic", "Biological", "Radiologicals", "Chemotherapeutic",
"Catalytics", "Hazardous", "Combustible", "Explosive", "Toxic", "Radioactive",
"Corrosive", "Flammable", "Reactive", "Stoichiometric", "Ionic", "Covalent",
"Thermodynamic", "Kinetic", "Aqueous", "Saturated", "Unsaturated",
"Volatile", "Nonvolatile", "Exothermic", "Endothermic", "Equilibrium",
"Isotopic", "Redox", "Precipitation", "Titration", "Soluble", "Insoluble",
"Catalyzed", "Inhibitor", "Catalytic", "Molar", "Empirical", "Molary",
"Hydrophobic", "Hydrophilic", "Hygroscopic", "Luminous", "Photovoltaic",
"Electrolytic", "Crystalline", "Amorphous", "Homogeneous",
"Heterogeneous", "Supersaturated", "Unsaturatedly", "Chemisorption",
"Adsorbent", "Coagulation", "Adsorption", "Desorption",
"Ionization", "Extraction", "Crystallization", "Chromatographic",
"Quantitative", "Qualitative", "Isotopics", "Sublimation",
"Precipitate", "Electrophilic", "Nucleophilic", "Condensation",
"Volatility", "Iontropic", "Biodegradable", "Thermophilic",
"Photoreactive", "Aromatic", "Aliphatic", "Elastomeric",
"Polymeric", "Amphoteric", "Hydrolytic", "Fluorescent", "Phosphorescent",
"Radioresistant", "Combustion", "Pyrolytic", "Aerobic",
"Anaerobic", "Oxidative", "Reductive", "Hydrosensitive", "Thermochemical",
"Crisis", "Endless")

```

## Comprobar número de elementos en los vectores

Comprobamos el número de elementos de ambos vectores. Como podemos ver, en el listado de nombres de laboratorios hay 151 elementos, uno más de los que pedimos. En este punto solo sabemos eso. ¿Podría haber un nombre repetido? ¿Cuál?

```
length(lab_names)
```

```
## [1] 151
```

```
length(lab_adjectives)
```

```
## [1] 150
```

## Buscar elementos únicos y duplicados

Nos aseguramos de que no hay ningún nombre o adjetivo repetido en las listas. Podemos utilizar la función **length** y **unique** para obtener los nombres de elementos ÚNICOS (no repetidos), y su número. Si nos fijamos bien, vemos que el número de elementos únicos que obtenemos en el vector `lab_names` es de 150, cuando sabemos ya que había 151. Esto se debe a que la función **unique** nos da los nombres **que no se repiten**, lo que indica que hay uno repetido.

Podemos saber cuales son los elementos repetidos con la función **duplicated**. Este ejemplo, el nombre “QuantumGen” está repetido, y habría que eliminarlo para que tener 150 nombres y que, además, que sean únicos.

```
length(unique(lab_names))
```

```
## [1] 150
```

```
length(unique(lab_adjectives))
```

```
## [1] 150
```

```
lab_names[duplicated(lab_names)]
```

```
## [1] "QuantumGen"
```

Para coger una muestra del vector nombres, y otra del de adjetivos, creamos primero dos vectores mucho más grandes, con 1.000 elementos seleccionados al azar en cada uno de ellos utilizando la función **sample**. Posteriormente tomamos **una** muestra de estos, y los guardamos en **n1** y **n2**. Finalmente los unimos en **n3** con la función **paste**. Con **print** se nos muestra en pantalla el nombre final de la empresa.

```
lab_names_1000 <- sample(lab_names, 1000, replace = TRUE)
lab_adjectives_1000 <- sample(lab_adjectives, 1000, replace = TRUE)
```

```
n1 <- sample(lab_names_1000, 1, replace = TRUE)
n2 <- sample(lab_adjectives_1000, 1, replace = TRUE)
```

```
n3 <-<- paste(n1, n2)
```

```
print(n3)
```

```
## [1] "GenoLab Veterinary"
```

## Función para generar $n$ nombres

Por último creamos la función `nombre_empresa`, con la que podremos generar el número de nombres que nos interese.

```
nombre_empresa <- function(numero) {

lab_names <- c("BioSynth", "ChemixLab", "PharmaWave", "LifeCore", "NanoBio", "GenoTechnics",
               "QuantumChem", "BioGene", "MediSynth", "LabPro", "Bioteched", "EnviroLab",
               "SynthCorp", "BioPharma", "GenoTech", "ChemWare", "MediGene", "QuantumBio",
               "LifeLab", "NanoTech", "SynthLife", "ChemixGen", "BioInno", "PharmaSyn",
               "GenoLab", "MediLife", "BiotechLab", "LabBio", "EnviroGen", "NanoPharm",
               "ChemCore", "Lifer", "QuantumCore", "SynthTech", "MediPharm", "GenoWave",
               "BioGenix", "LabSynth", "PharmaCorp", "NanoCore", "ChemBio", "EnviroTech",
               "LifeTech", "QuantumGen", "MediGenix", "BioCorp", "GenoBio", "SynthWave",
               "BiotechX", "LabGen", "NanoLife", "PharmaInno", "EnviroBio", "ChemPro",
               "LifeInno", "QuantumPharm", "MediPro", "GenoInno", "BioTech", "SynthPharm",
               "LabPharm", "NanoSynth", "LifeBio", "EnviroSynth", "MediSynthX", "GenoPro",
               "ChemSynth", "BioPharmX", "QuantumLab", "NanoPharma", "LifePharm", "SynthInno",
               "MediPharma", "GenoCorp", "BioWare", "LabLife", "EnviroInno", "NanoGenix",
               "ChemProX", "QuantumBios", "MediCore", "SynthLifeX", "BioLifeX", "PharmaTech",
               "LifeGenix", "GenoPharmaX", "NanoBioX", "ChemTechX", "MediBioX", "SynthPharmaX",
               "LabGenix", "EnviroGenix", "NanoGen", "BioGen", "LifeProX", "MediPharmX", "GenoGenixX",
               "BioSolve", "ChemBioX", "GenoNova", "NanoCoreX", "SynthPharmaZ", "LabInnoX",
               "EnviroLifeZ", "MediTechX", "QuantumProZ", "LifeGenius", "PharmaBioZ",
               "BioTechLabZ", "GenoScienceZ", "NanoPharmX", "ChemXcelZ", "SynthoMedX",
               "MediWaveZ", "LabPulseX", "EnviroTechZ", "GenoSynthX", "BioMaxZ",
               "NanoBioTechZ", "LifeTechX", "QuantumBioX", "ChemWaveZ", "SynthoLifeX",
               "LabCoreZ", "GenoTechX", "EnviroPharmX", "BioInnoZ", "NanoProX", "LifePharmZ",
               "PharmaGenX", "QuantumCoreX", "MediSynthZ", "ChemXpertX", "SynthoGenX",
               "LabXpressZ", "GenoInnoZ", "EnviroBioX", "BioTechXcelX", "NanoPharmaZ",
               "LifeGenixX", "PharmaTechZ", "QuantumBioXcelX", "MediWaveXcelX", "ChemCoreXcelZ",
               "SynthoPharmaX", "LabTechZ", "BioGenXcel", "NanoCoreInno", "LifePharmTech",
               "SynthoMedInnovations", "QuantumGen")

lab_adjetives <- c("Innovative", "Technological", "Biologicals", "Therapeutic", "Pharmacological",
                   "Biogenetic", "Clinical", "Modern", "Efficient", "Investigative",
                   "Experimental", "Biomedical", "Molecular", "Engineering",
                   "Precise", "Analyticals", "Synthetics", "Futuristic", "Personalized",
                   "Specialized", "Genetic", "Nano", "Neuropharmaceutical",
                   "Ecological", "Veterinary", "International", "Scientific",
                   "Inspiring", "Sustainable", "Advanced", "Revolutionary", "Healthy",
                   "Radiological", "Biochemicals", "Cloning", "Environmentaly", "Regenerative",
                   "Bioinformatics", "Microscopic", "Pioneering", "Metabolic",
                   "Pharmaceuticals", "Medical", "Intelligent", "Genomic", "Chemical", "Analytical",
                   "Organic", "Inorganic", "Biochemical", "Physical", "Nuclear", "Synthetic",
                   "Molecular", "Polymer", "Atomic", "Pharmaceutical", "Environmental",
                   "Quantum", "Spectroscopic", "Biological", "Radiologicals", "Chemotherapeutic",
                   "Catalytics", "Hazardous", "Combustible", "Explosive", "Toxic", "Radioactive",
                   "Corrosive", "Flammable", "Reactive", "Stoichiometric", "Ionic", "Covalent",
                   "Thermodynamic", "Kinetic", "Aqueous", "Saturated", "Unsaturated",
                   "Volatile", "Nonvolatile", "Exothermic", "Endothermic", "Equilibrium",
                   "Isotopic", "Redox", "Precipitation", "Titration", "Soluble", "Insoluble",
                   "Catalyzed", "Inhibitor", "Catalytic", "Molar", "Empirical", "Molarity",
```

```

"Hydrophobic", "Hydrophilic", "Hygroscopic", "Luminous", "Photovoltaic",
"Electrolytic", "Crystalline", "Amorphous", "Homogeneous",
"Heterogeneous", "Supersaturated", "Unsaturatedly", "Chemisorption",
"Adsorbent", "Coagulation", "Adsorption", "Desorption",
"Ionization", "Extraction", "Crystallization", "Chromatographic",
"Quantitative", "Qualitative", "Isotopics", "Sublimation",
"Precipitate", "Electrophilic", "Nucleophilic", "Condensation",
"Volatility", "Ionotropic", "Biodegradable", "Thermophilic",
"Photoreactive", "Aromatic", "Aliphatic", "Elastomeric",
"Polymeric", "Amphoteric", "Hydrolytic", "Fluorescent", "Phosphorescent",
"Radioresistant", "Combustion", "Pyrolytic", "Aerobic",
"Anaerobic", "Oxidative", "Reductive", "Hydrosensitive", "Thermochemical",
"Crisis", "Endless")

lab_names_1000 <- sample(lab_names, 1000, replace = TRUE)
lab_adjectives_1000 <- sample(lab_adjectives, 1000, replace = TRUE)

n1 <- sample(lab_names_1000, numero, replace = TRUE)
n2 <- sample(lab_adjectives_1000, numero, replace = TRUE)
n3 <-<- paste(n1, n2)

return(n3)
}

```

Si, por ejemplo, queremos que nos ofrezca **10 nombres**, llamamos a la función de la siguiente forma. Los nombres resultantes de las empresas serán listados.

```
nombre_empresa(10)
```

```
## [1] "ChemXpertX Pharmaceutical" "LabInnoX Hygroscopic"
## [3] "MediBioX Chemotherapeutic" "BioSolve Crystallization"
## [5] "PharmaWave Biologicals" "GenoSynthX Chemotherapeutic"
## [7] "ChemixGen Coagulation" "SynthoLifeX Toxic"
## [9] "GenoBio Redox" "GenoTechnics Thermochemical"
```