

Choosing the Right Data Protection Strategies For Your Ceph Deployments

Ceph Day Seattle 2025
Anthony Datri (IBM)



Choosing data protection strategy can be complicated

- 🐙 Usable to raw capacity ratio
- 🐙 Replication vs erasure coding
- 🐙 EC profile values for k and m
- 🐙 Read and Write performance
- 🐙 Recovery performance
- 🐙 Failure domains
- 🐙 Fault tolerance
- 🐙 Media saturation
- 🐙 `min_alloc_size` vs IU

Replication vs Erasure Coding



Replication

- Three (usually) copies of all data



Erasure Coding (EC)

- Splits data into chunks
- Computes and stores additional parity chunks
- Akin to RAID6 / RAID60 but more flexible

Replication



- Fast and simple

- As few as 3 nodes or racks

- High space amplification

- Ratio of *raw* capacity to *usable* capacity
- Default 3 copies: space amp factor of **3**
- Other values are possible: 1 and 2 are dangerous if your data isn't disposable
- Stretch mode clusters require 4



Erasure Coding (EC)

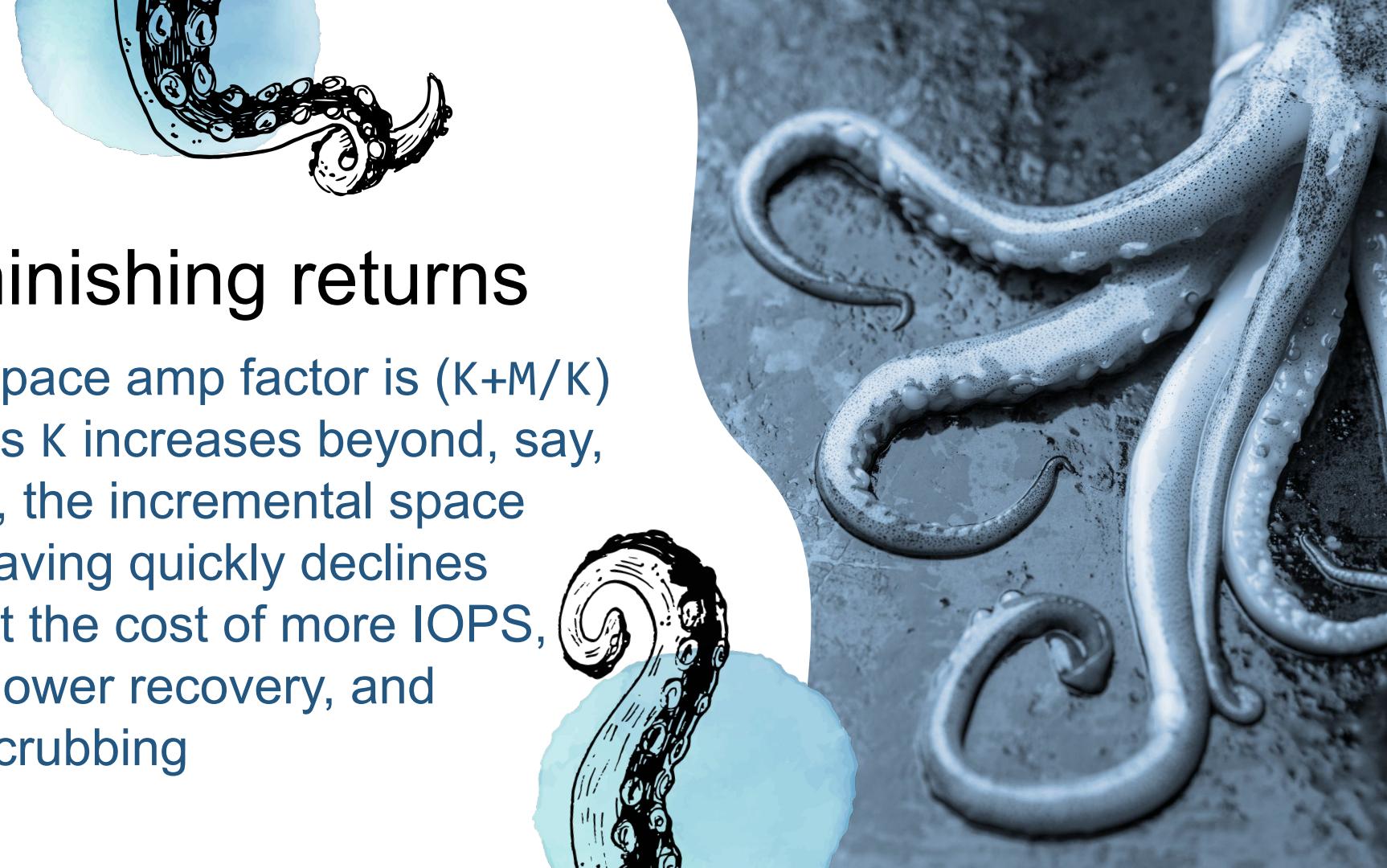
- Octopus icon: Requires additional CPU for parity calculation
- Octopus icon: Usually requires more hosts/racks than Replication
- Octopus icon: Less space amp
 - EC is often selected to maximize *usable* capacity for a given amount of *raw* underlying storage
 - Or to minimize *raw* capacity for desired *usable*
 - More TB/RU, TB/node, TB/watt
- Octopus icon: Extra IOPs on underlying drives: HDDs saturate
- Octopus icon: Slower recovery; client writes may have higher latency
- Octopus icon: May burn SSD endurance more quickly (though this is mostly FUD)(mostly)

EC Profile

- EC pools have attributes: most important are K and M
- K is the number of data chunks: when K=4, 1MB of user data is split into 4 data chunks @ 256KB
- M is the number of coding chunks, which are the same size as data chunks
- When M=2, 1MB of user data generates 2 coding chunks @ 256KB
- We call the above a 4+2 or 4, 2 profile
- Space amp factor here is just 1.5

Diminishing returns

- Space amp factor is $(K+M/K)$
- As K increases beyond, say, 6, the incremental space saving quickly declines
- At the cost of more IOPS, slower recovery, and scrubbing



Erasure coding overhead

	m=1	m=2	m=3	m=4	m=5	m=6	m=7	m=8	m=9	m=10	m=11
k=1	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00
k=2	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50
k=3	1.33	1.67	2.00	2.33	2.67	3.00	3.33	3.67	4.00	4.33	4.67
k=4	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75
k=5	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00	3.20
k=6	1.16	1.33	1.50	1.66	1.83	2.00	2.17	2.33	2.50	2.66	2.83
k=7	1.14	1.29	1.43	1.58	1.71	1.86	2.00	2.14	2.29	2.43	2.58
k=8	1.13	1.25	1.38	1.50	1.63	1.75	1.88	2.00	2.13	2.25	2.38
k=9	1.11	1.22	1.33	1.44	1.56	1.67	1.78	1.88	2.00	2.11	2.22
k=10	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	2.00	2.10
k=11	1.09	1.18	1.27	1.36	1.45	1.54	1.63	1.72	1.82	1.91	2.00
k=12	1.08	1.17	1.25	1.33	1.42	1.50	1.58	1.67	1.75	1.83	1.92
k=20	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	1.50	1.55



Don't. Just....Don't.

High risk of data loss
Data is unavailable if even
one host / rack is down



Selecting K



Tradeoff of space amp, performance, fault tolerance



Data is *preserved* if any K data **or** coding chunks survive



Data is *available* for reads and writes if any K+1 data or coding chunks are online



Dan's presentation will explore the durability and availability benefits of larger values of K



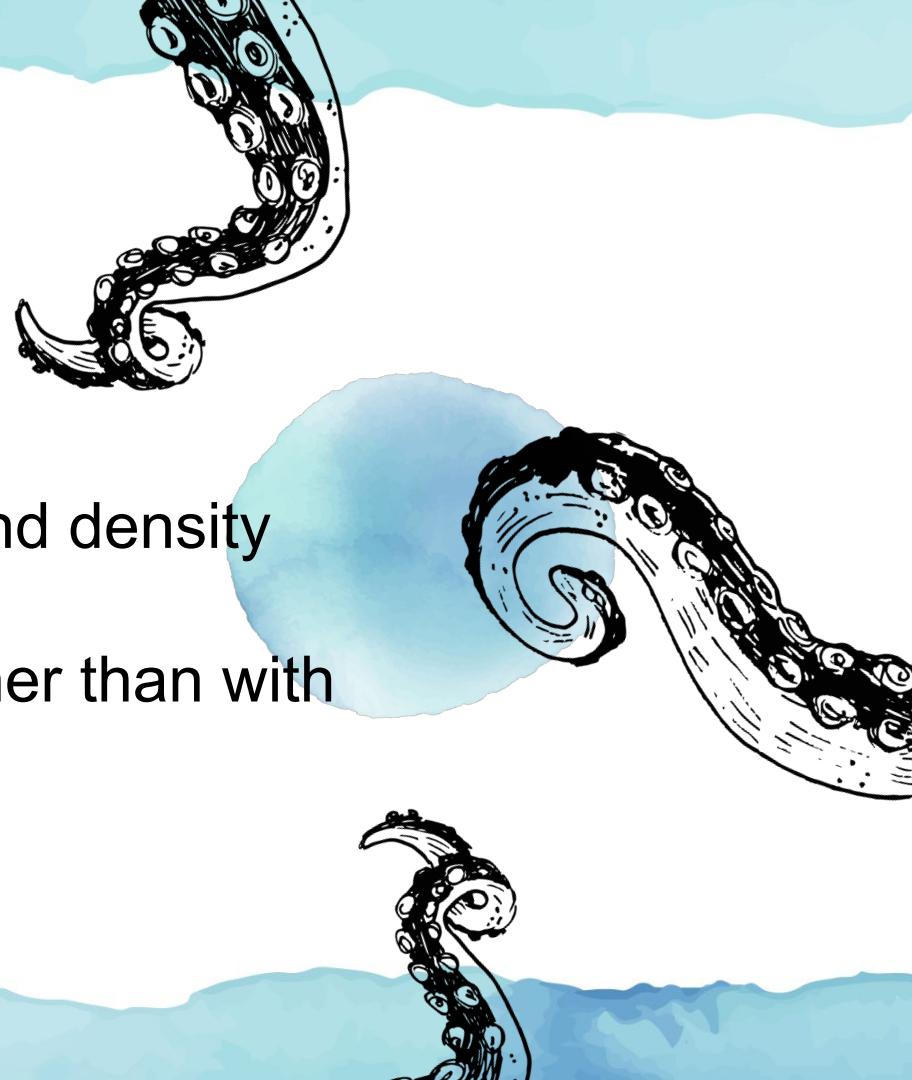
3+3, 6+3 are examples of profiles with higher fault tolerance at the expense of performance and space amp

When to choose Replication or EC

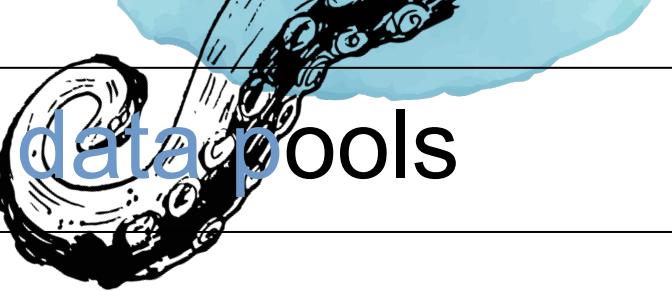
- EC is best across at least $M+K+1$ failure domains
- You *can* do EC on fewer, but it's tricky
- Configured at pool creation
- Can't switch between the two later
- Some pools **require** replication:
 - RGW index, CephFS metadata
- RBD pools are almost always replicated
- RBD using EC is possible since Kraken
 - Adjunct replicated metadata pool
 - Higher latency is usually unacceptable

EC for CephFS and RGW data pools

- Applications are often less sensitive to latency
- Driven by space amp: cost and density over performance
- Throughput *can* even be higher than with replication



EC for CephFS and RGW **data pools**



-  You can do both!
 -  Multiple CephFS data pools tied to directories
 -  Multiple RGW storage classes and placement targets
 - Clients can select via HTTP header
 - Users and buckets can have a default set
 - Lua scripting can override header/default
 - RFE submitted to direct data to pools by policy
- 
- 

Multiple RGW data pools



- EC data pools can result in larger space amp for very small objects (this may be improved in Tentacle)
- Small / hot RGW objects benefit from replicated pool perf
- User objects in a secondary data pool still have a HEAD RADOS object in the default; some ops benefit from speed
- Larger / tepid / cold objects often are fine on an EC pool for efficiency



Mixed media RGW + CephFS data pools



Best of both worlds

Fast TLC SSDs for replicated metadata + tiny / hot objects

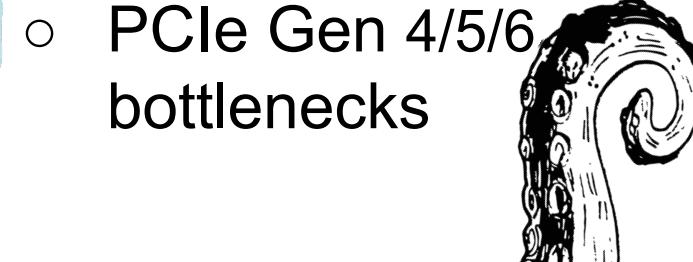
- Possible inlining of small user objects / files

Dense, cost-effective HDD for EC data pools

- Spinners gonna seek

Denser pTLC or QLC SSDs for EC data pools

- Up to 122TB, larger on the horizon
- PCIe Gen 4/5/6 allow huge capacity without bottlenecks



Reference



Mark Nelson's space amp cheat sheet

<https://docs.google.com/spreadsheets/d/1rpGfScgG-GLoIGMJWDixEkqs-0n9w8nAUToP0jN8bDI/edit?usp=sharing>



EC nuts and bolts and overhead table

<https://docs.ceph.com/en/latest/rados/operations/erasure-code>



Anthony D'Atri

<https://www.linkedin.com/in/anthonydatri/>



THANKS

Do you have any questions?

youremail@freepik.com

+91 620 421 838

yourcompany.com



CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#) and infographics & images by [Freepik](#)

Please keep this slide for attribution





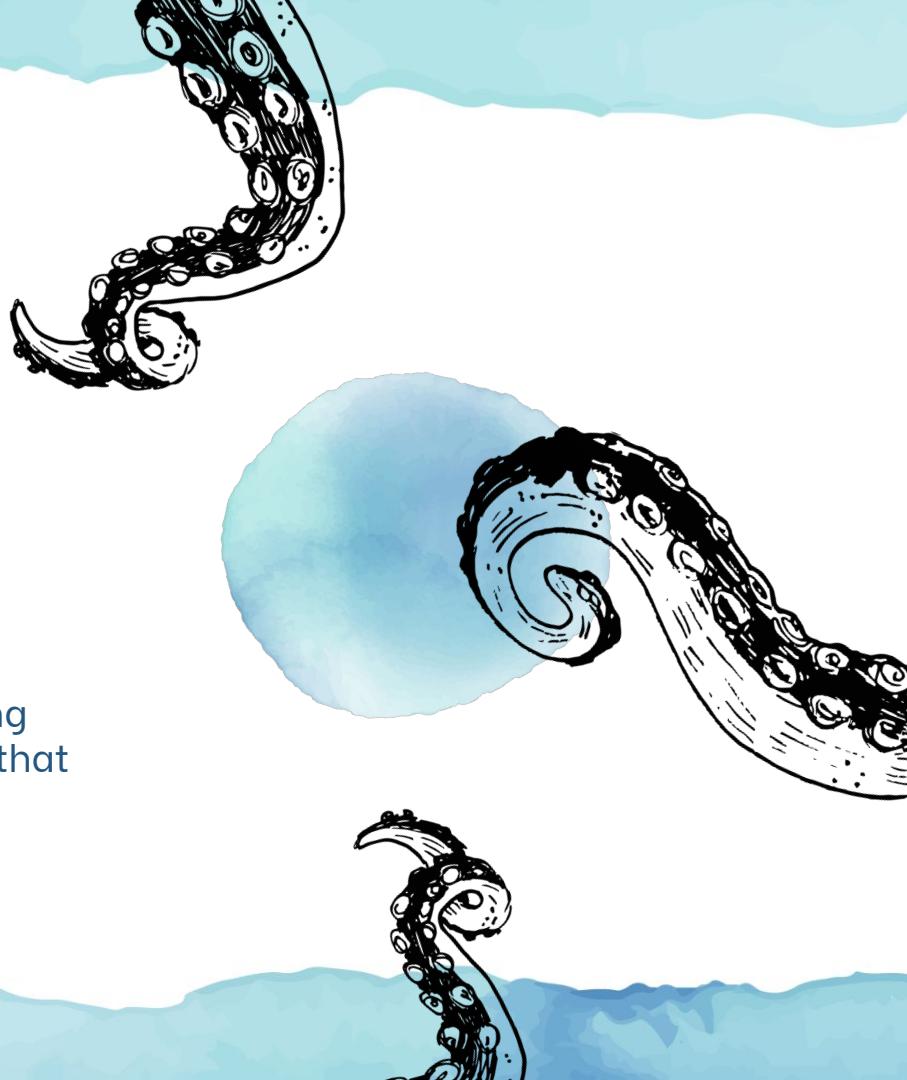
May the Tentacle be with you

(friends not food)



WHOA!

Octopus fossils have been discovered dating back over 300 million years. This indicates that the octopus predates the dinosaur!



May the
Tentacle be with
you



CONTENTS OF THIS TEMPLATE

Here's what you'll find in this **Slidesgo** template:

1. A slide structure based on a multipurpose presentation, which you can easily adapt to your needs. For more info on how to edit the template, please visit **Slidesgo School** or read our **FAQs**.
2. An assortment of graphic resources that are suitable for use in the presentation can be found in the **alternative resources slide**.
3. A **thanks slide**, which you must keep so that proper credits for our design are given.
4. A **resources slide**, where you'll find links to all the elements used in the template.
5. **Instructions for use**.
6. Final slides with:
 - The **fonts and colors** used in the template.
 - A selection of **illustrations**. You can also customize and animate them as you wish with the online editor. Visit **Storyset** to find more.
 - More **infographic resources**, whose size and color can be edited.
 - Sets of **customizable icons** of the following themes: general, business, avatar, creative process, education, help & support, medical, nature, performing arts, SEO & marketing, and teamwork.

You can delete this slide when you're done editing the presentation

TABLE OF CONTENTS

01

ABOUT THE DAY

You can describe the topic of the section here

03

FUTURE PLANS

You can describe the topic of the section here

02

DATA ANALYTICS

You can describe the topic of the section here

04

SOCIAL MEDIA

You can describe the topic of the section here

INTRODUCTION

October 8th is when we celebrate this day. Octopuses are marine organisms with eight long arms, spherical bodies, and projecting eyes. Because they reside in deep seas, they are known as “monsters of the deep”





01

ABOUT THE DAY

You can enter a subtitle here if you need it

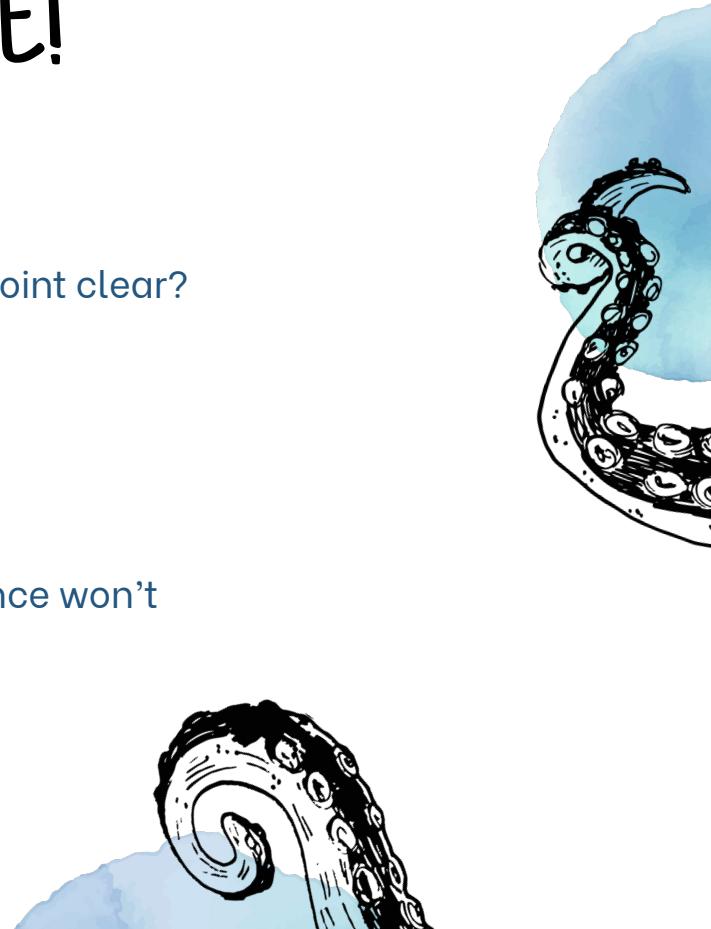


THE SLIDE TITLE GOES HERE!

Do you know what helps you make your point clear?
Lists like this one:

- They're simple
- You can organize your ideas clearly
- You'll never forget to buy milk!

And the most important thing: the audience won't miss the point of your presentation



IMPORTANT INFO



BRAIN

Octopuses have decentralized brains that are spread all over their eight limbs



VENOM

An octopus can catch its victim, pierce the shell and inject its toxic saliva into it

DID YOU KNOW?



PREDATORS

Almost all octopuses are predatory eating a large variety of species



REGROWTH

The regrowth process will start as soon as they lose their tentacle



SMART

The problem-solving skills are evidence of the evolution of its intelligence



HOW TO FIND THEM



TRIP 1

Saturn is a gas giant and has several rings



TRIP 2

Jupiter is the biggest planet in the Solar System



TRIP 3

Despite being red, Mars is actually a cold place



TRIP 4

Venus is the second planet from the Sun



SOME SPECIES

BLUE RINGED

Mercury is the closest planet to the Sun

MIMIC

Venus is the second planet from the Sun

COMMON

Mars is actually a very cold place

COCONUT

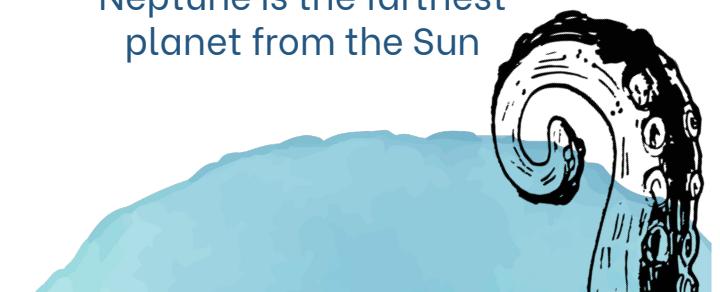
Jupiter is the biggest planet of them all

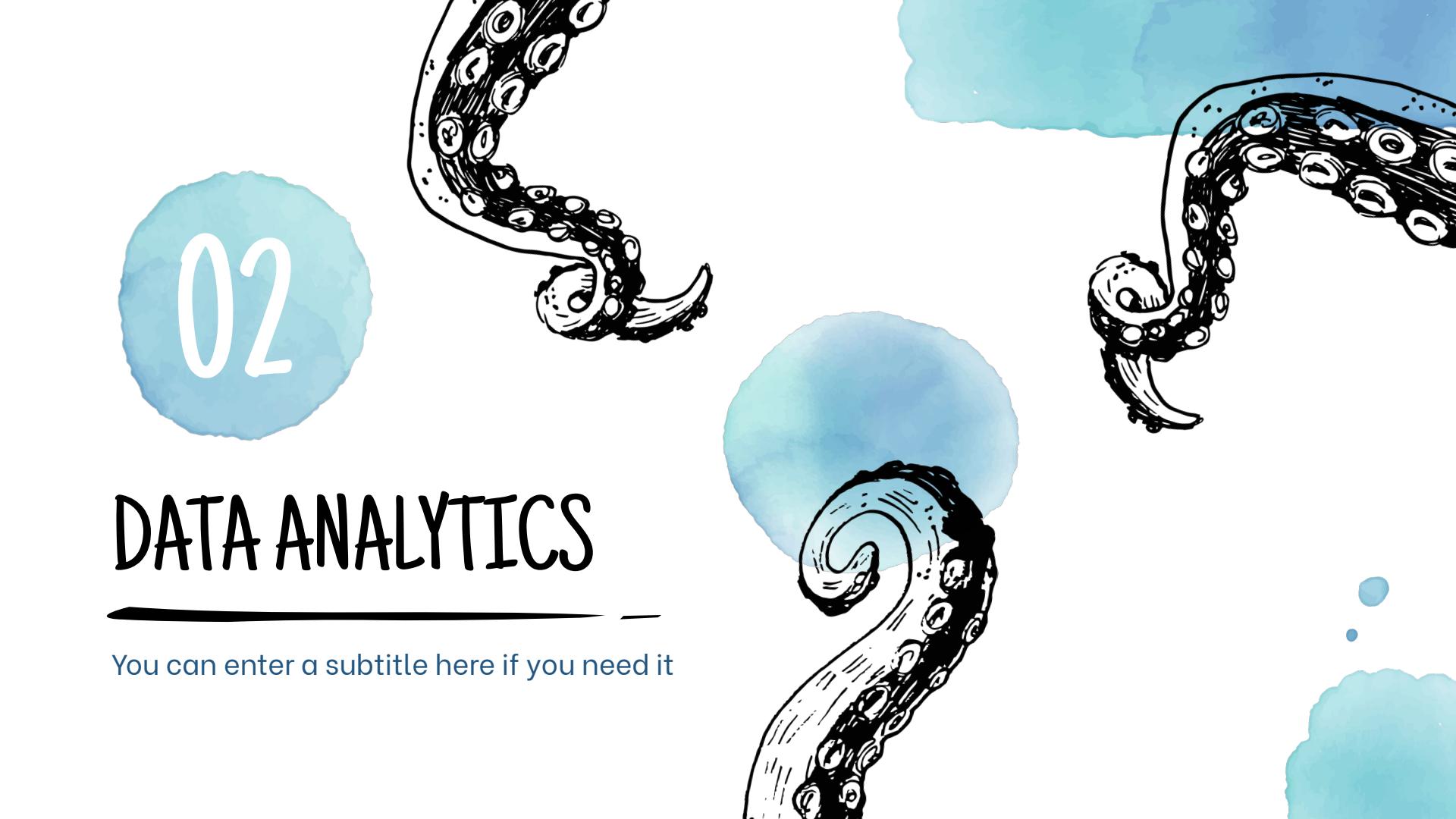
BLANKET

Saturn is composed of hydrogen and helium

CARIBBEAN

Neptune is the farthest planet from the Sun





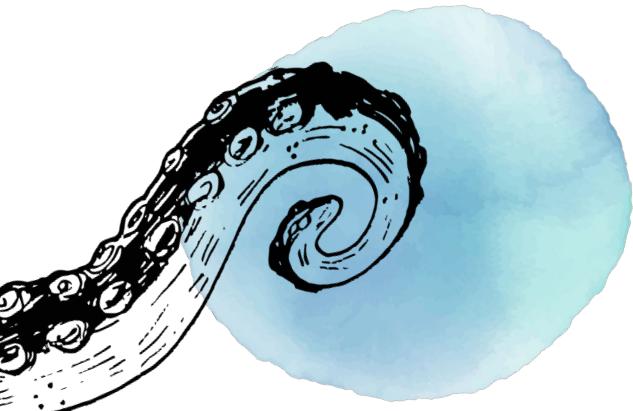
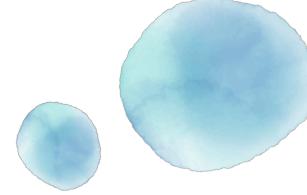
02

DATA ANALYTICS

You can enter a subtitle here if you need it

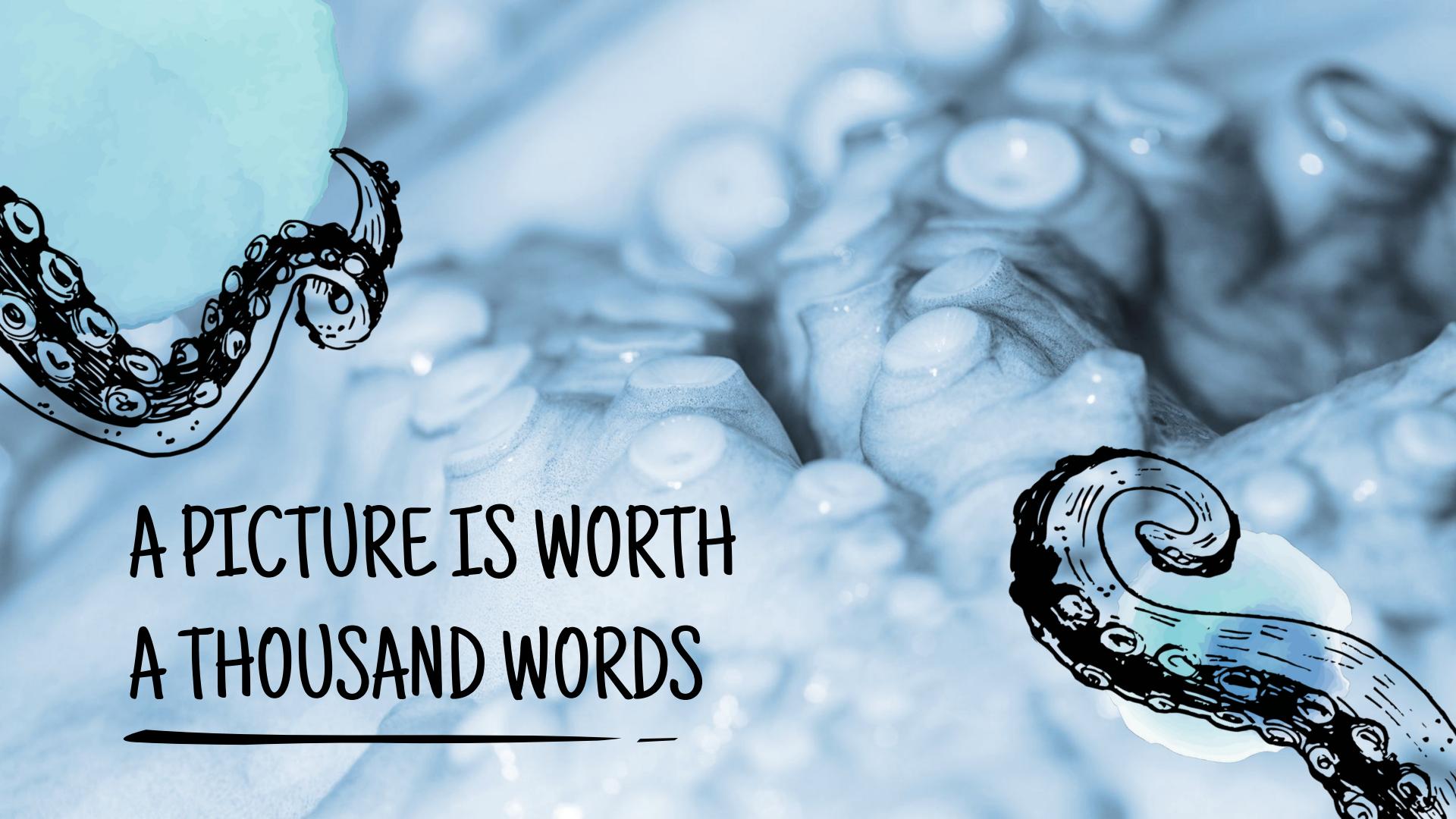
AWESOME WORDS





“This is a quote, words full of wisdom
that someone important said and
can make the reader get inspired.”

—SOMEONE FAMOUS



A PICTURE IS WORTH
A THOUSAND WORDS





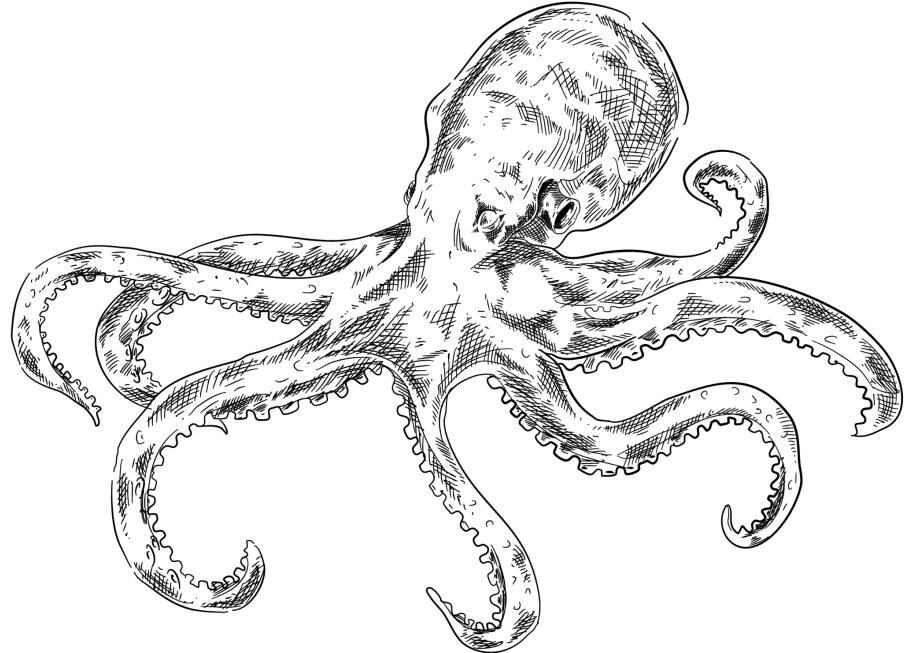
A PICTURE ALWAYS REINFORCES THE CONCEPT



Images reveal large amounts of data, so remember: use an image instead of a long text. Your audience will appreciate it

300

Are the different species of octopuses



500M

Neurons in their brains
and arms

156

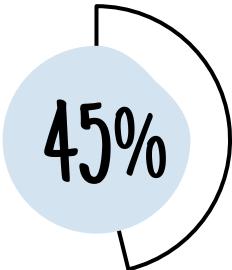
Pounds was the record
weight for them

25

Miles per hour is their
maximum speed



SOME PERCENTAGES



MERCURY

Mercury is the closest planet to the Sun



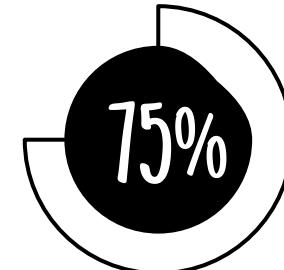
VENUS

Venus is the second planet from the Sun



JUPITER

Jupiter is the biggest planet of them all



EARTH

The Earth is the planet where we all live



A black and white line drawing of several tentacles with suckers, some coiled and some straight, set against a background of light blue watercolor washes in circular and organic shapes.

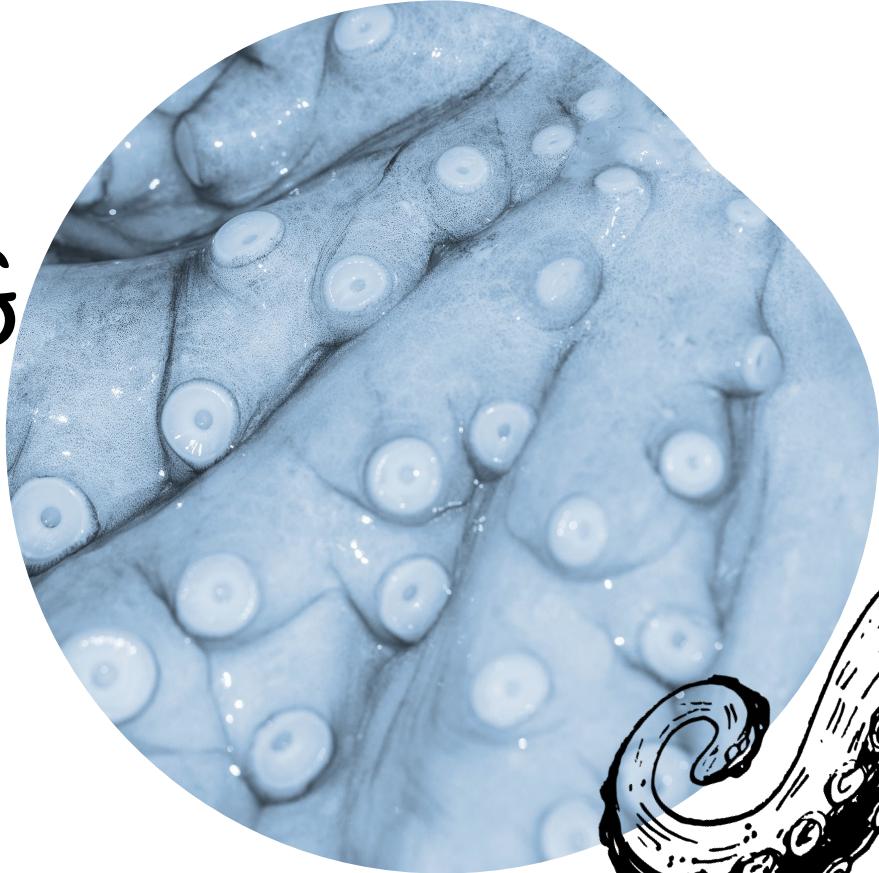
03

FUTURE PLANS

You can enter a subtitle here if you need it

INSERT AN AMAZING TITLE HERE

You also have space for an
incredible subtitle!



BEST TRIPS TO FIND THEM!

\$160

BASIC PACK

Venus is the second planet from the Sun

\$180

PREMIUM PACK

Saturn is made of hydrogen and helium

\$199

ULTIMATE PACK

Jupiter is the biggest planet of them all



THIS TEXT SEEKS TO BE IMPORTANT...

The octopus is compared to Houdini, in addition to being extremely intelligent. Because the octopus lacks an internal skeleton or a protective shell, it may squeeze through small gaps. An octopus has been reported to escape from a jar with a closed cover due to its lack of bone structure. They are truly one of a kind and extraordinary creatures!



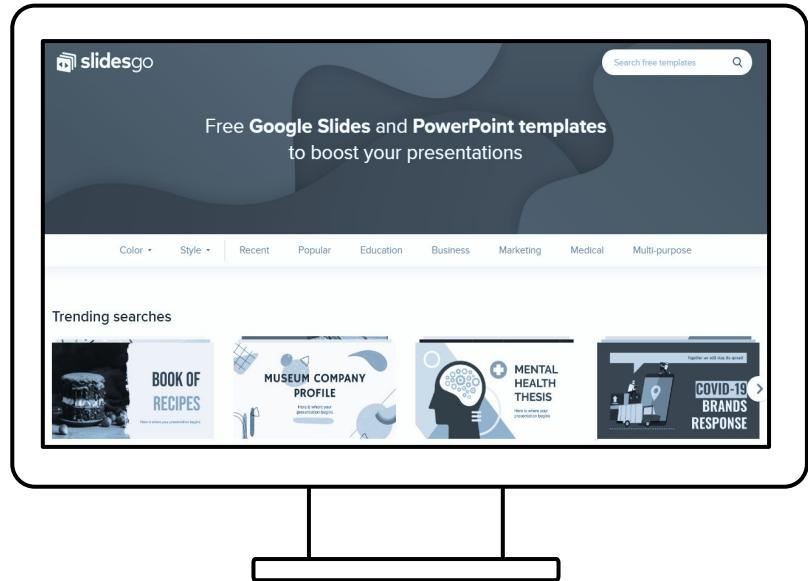
04

SOCIAL MEDIA

You can enter a subtitle here if you need it

DESKTOP SOFTWARE

You can replace the image on the screen with your own work. Just right-click on it and select “Replace image”



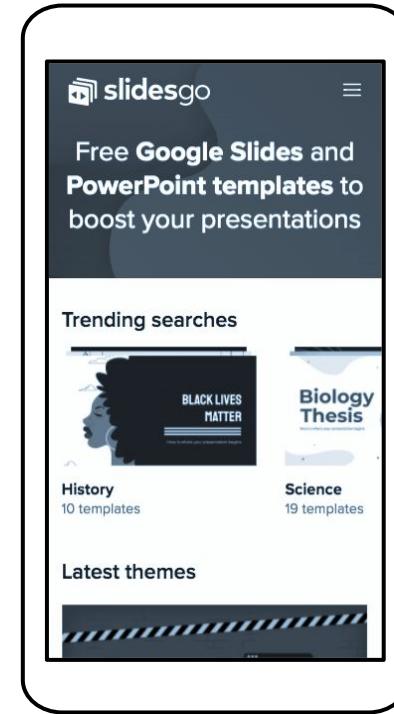


TABLET APP

You can replace the image on the screen with your own work. Just right-click on it and select “Replace image”

MOBILE APP

You can replace the image on the screen with your own work. Just right-click on it and select “Replace image”



KNOW OUR TEAM!



JAMES DOE

You can speak a bit about
this person here



TIM JIMMY

You can speak a bit about
this person here



JOHN BONES

You can speak a bit about
this person here

BEST LOCATIONS TO FIND THEM

DEEP SEA

Venus is the second planet from the Sun

OPEN SEA

Despite being red, Mars is a cold place



SCHEDULE FOR NEXT TRIPS



Mercury is closest planet to the Sun and the smallest



Despite being red, Mars is actually a very cold place



AUG

SEP

OCT

NOV

DIC

Venus has a beautiful name and high temperatures



Neptune is extremely far from the Sun



Saturn is the only gas giant that has incredible rings



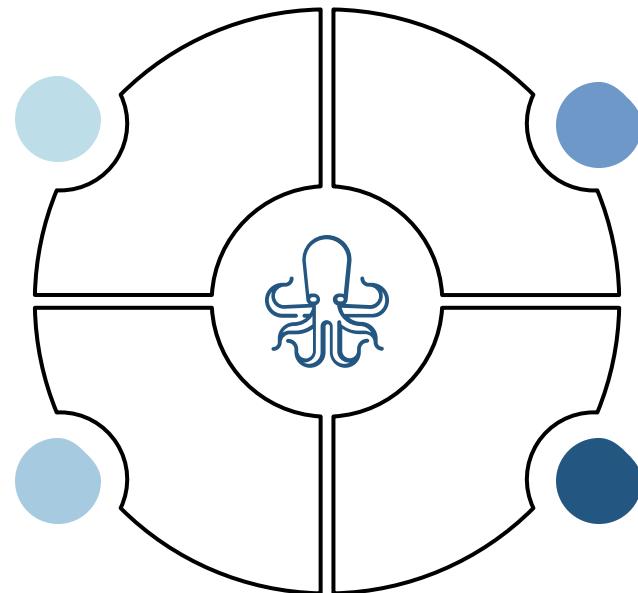
INTERESTING FACTS

BLOOD

Octopuses' blood is extrangely bluish

BREATH

Octopuses breathe through the gills



PROTEIN

Octopus are rich in the protein hemocyanin

COLOR

Octopuses can change colors in an instant



SPECIES CHARACTERISTICS

	STRENGTH	INTELLIGENCE	SIZE
BLANKET	9/10	7/10	5/10
COMMON	7/10	8/10	7/10
MIMIC	6/10	9/10	6/10

INTERESTING FACTS

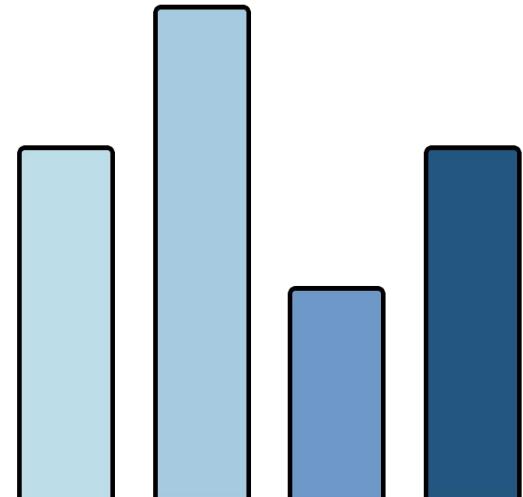
40

30

20

10

0



25%

JUPITER

It's the biggest planet
in the Solar System

15%

SATURN

Saturn is a gas giant
and has several rings

35%

VENUS

Venus is the second
planet from the Sun

15%

MARS

Despite being red, Mars
is a cold place

Follow the link in the graph to modify its data and then paste the new one here. [For more info, click here](#)



CONCLUSIONS

VENUS



Venus is the second planet from the Sun

MERCURY



Mercury is the closest planet to the Sun



SATURN



It's composed of hydrogen and helium

JUPITER



Jupiter is the biggest planet of them all



ALTERNATIVE RESOURCES

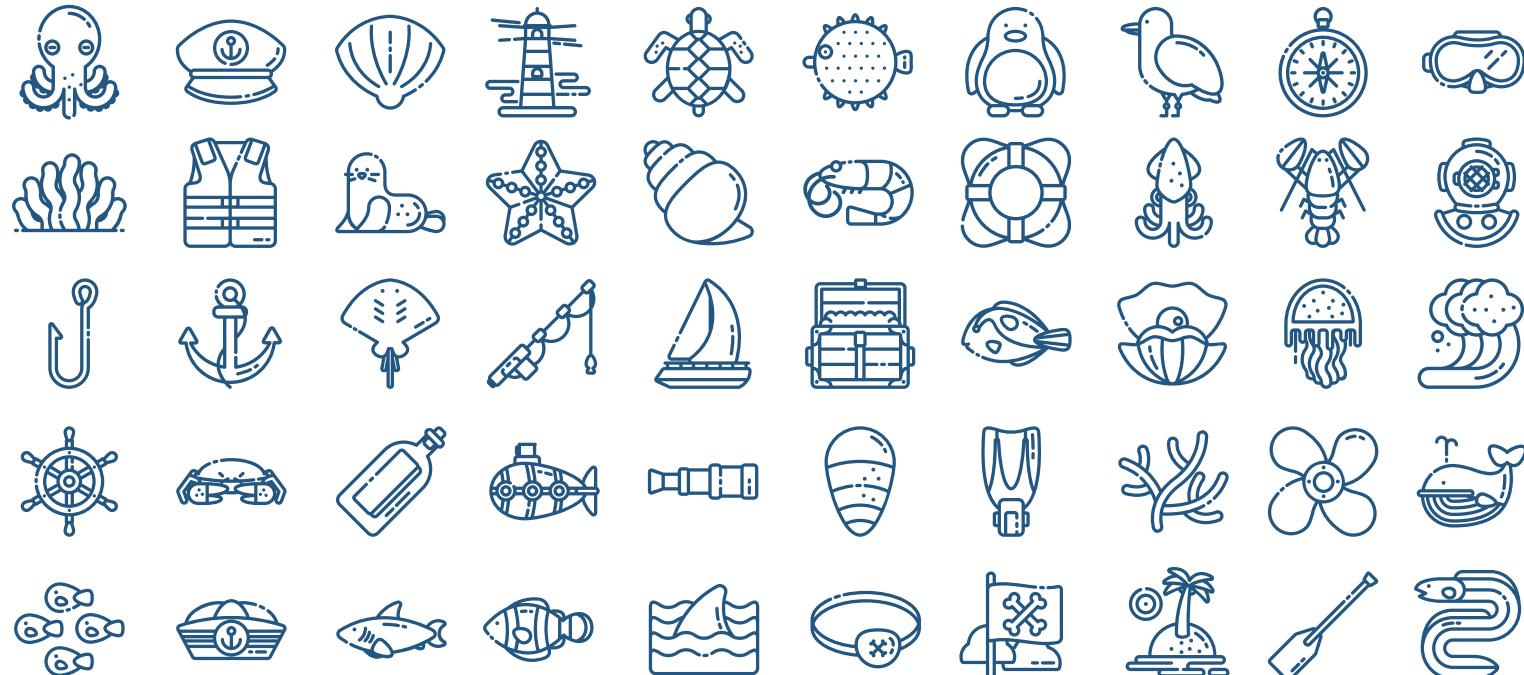
Here's an assortment of alternative resources whose style fits the one of this template

PHOTOS

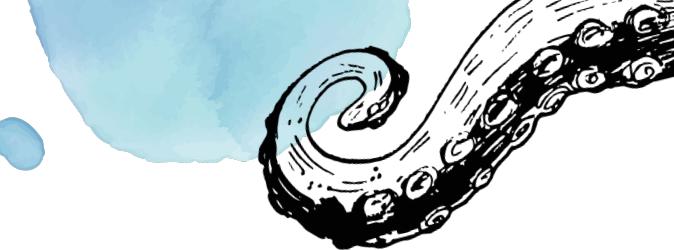
- Close-up delicious octopus on table
- Top view two octopus in opposite corners
- Copy-space with octopus on table
- Top view of squid on slate
- High angle of black pasta with squid
- Top view of squid on plate
- Close-up octopus with copy-space



ICON PACK: SEALIFE



RESOURCES



Did you like the resources on this template? Get them for free at our other websites:

VECTORS

- Abstract watercolor instagram posts
- Abstract watercolor instagram posts II
- Hand drawn restaurant business card
- Hand drawn design octopus

ICONS

- Icon Pack: Sealife | Lineal

PHOTOS

- Close-up exotic octopus seafood dish
- Raw octopus close-up wallpaper
- Front view fresh octopus on table
- Close-up of fresh octopus in market
- Close-up of octopus for sale
- People collage design

Instructions for use

In order to use this template, you must credit **Slidesgo** by keeping the **Thanks** slide.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>

Instructions for use (premium users)

In order to use this template, you must be a Premium user on [Slidesgo](#).

You are allowed to:

- Modify this template.
- Use it for both personal and commercial purposes.
- Hide or delete the “Thanks” slide and the mention to Slidesgo in the credits.
- Share this template in an editable format with people who are not part of your team.

You are not allowed to:

- Sublicense, sell or rent this Slidesgo Template (or a modified version of this Slidesgo Template).
- Distribute this Slidesgo Template (or a modified version of this Slidesgo Template) or include it in a database or in any other product or service that offers downloadable images, icons or presentations that may be subject to distribution or resale.
- Use any of the elements that are part of this Slidesgo Template in an isolated and separated way from this Template.
- Register any of the elements that are part of this template as a trademark or logo, or register it as a work in an intellectual property registry or similar.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>

Fonts & colors used

This presentation has been made using the following fonts:

Just Another Hand

(<https://fonts.google.com/specimen/Just+Another+Hand>)

Be Vietnam

(<https://fonts.google.com/specimen/Be+Vietnam>)

#000000

#235680

#6d98c7

#a6cbe0

#bdddee8

#d4e3f0

#ffffff

Storyset

Create your Story with our illustrated concepts. Choose the style you like the most, edit its colors, pick the background and layers you want to show and bring them to life with the animator panel! It will boost your presentation. Check out [How it Works](#).



Pana



Amico



Bro



Rafiki



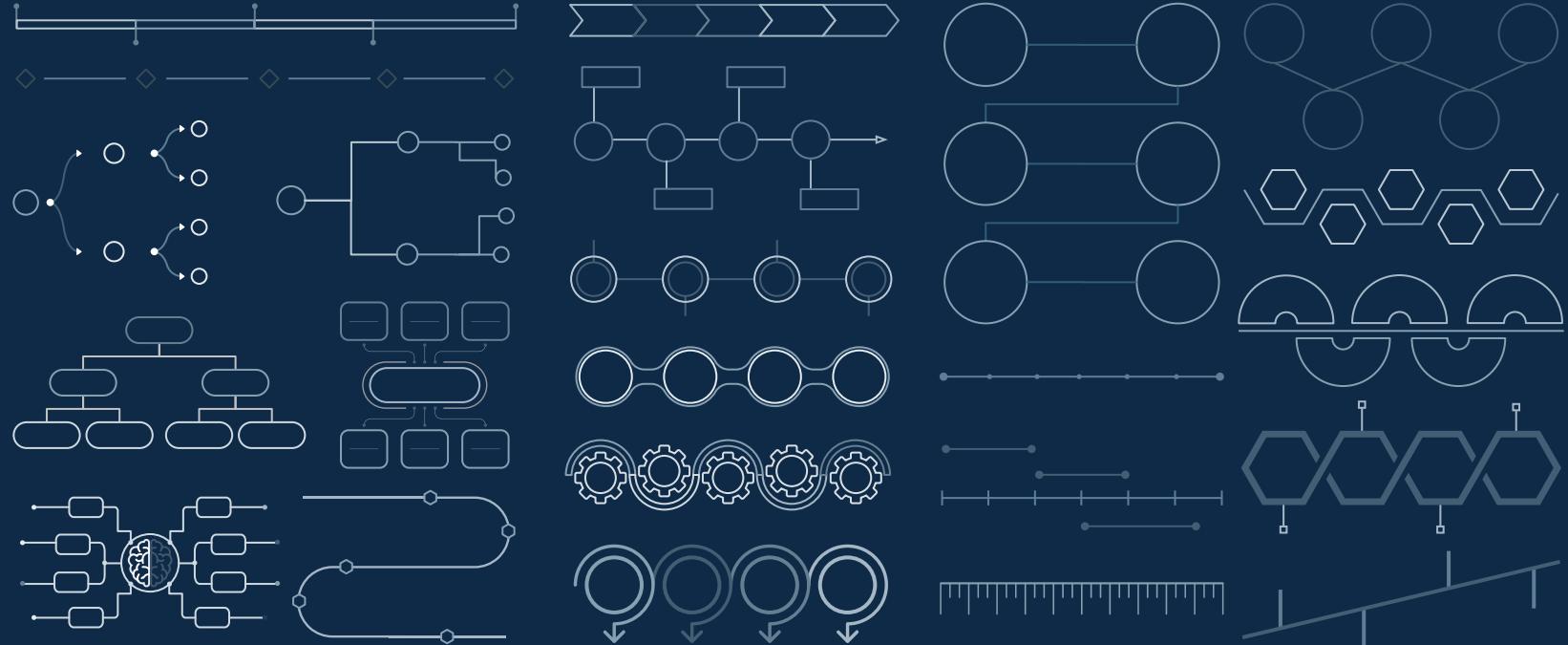
Cuate

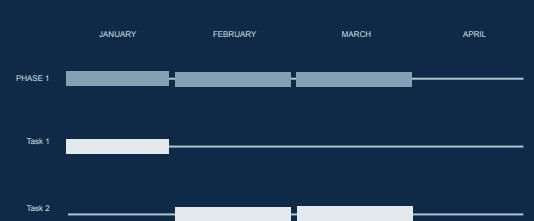
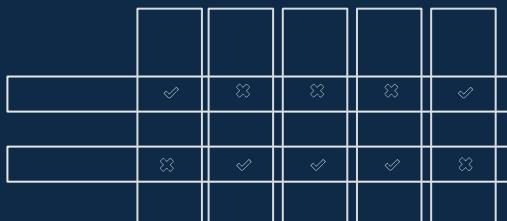
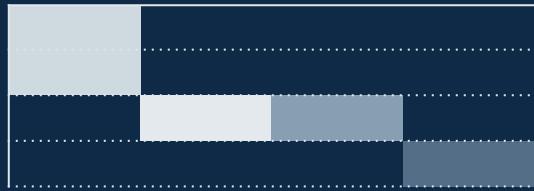
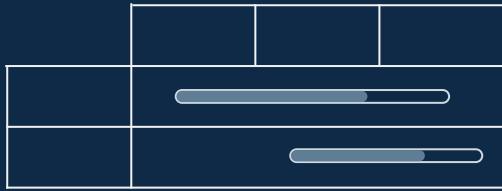
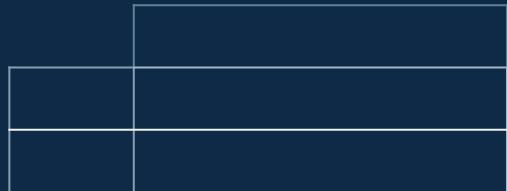
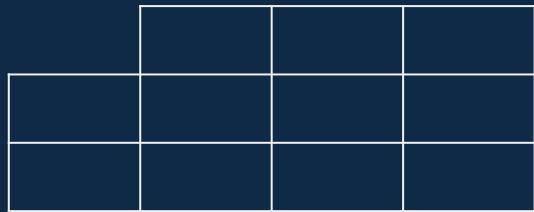
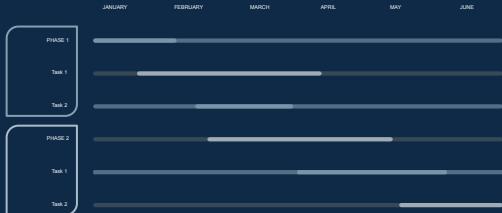
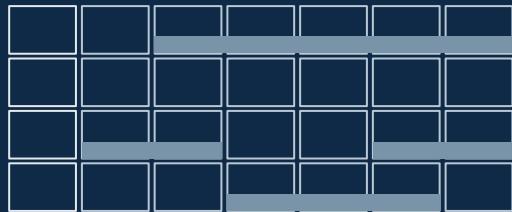
Use our editable graphic resources...

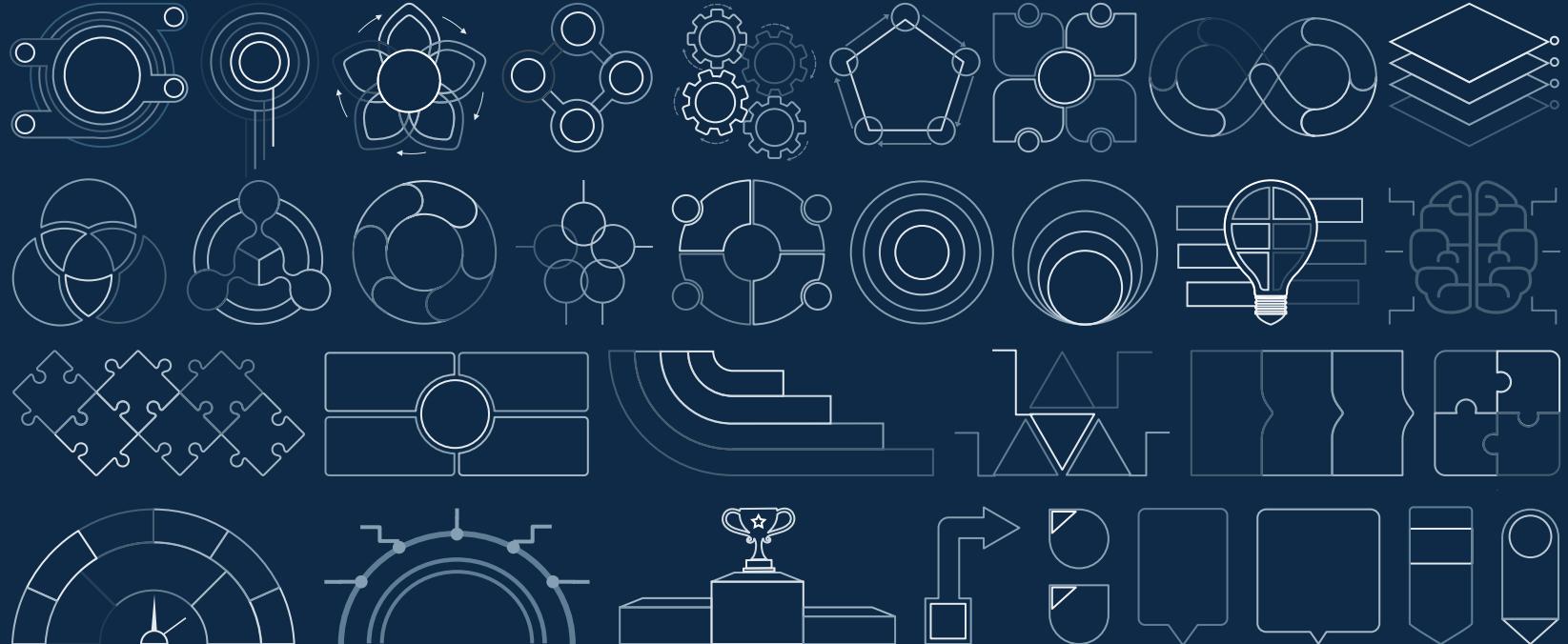
You can easily resize these resources without losing quality. To change the color, just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want. Group the resource again when you're done. You can also look for more infographics on Slidesgo.

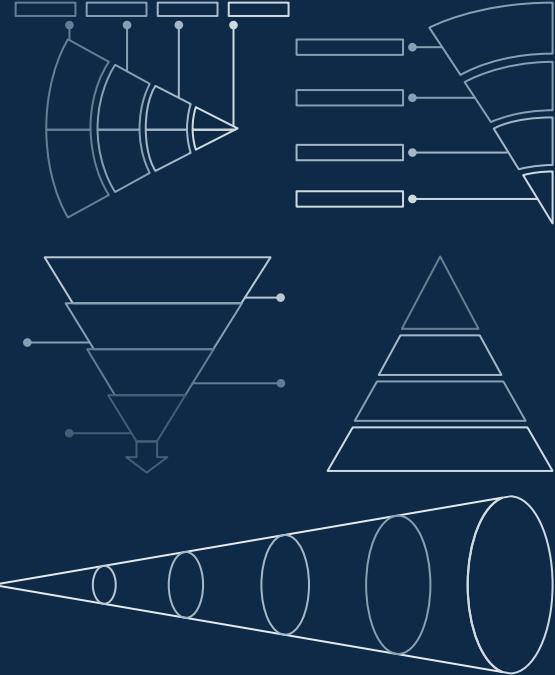
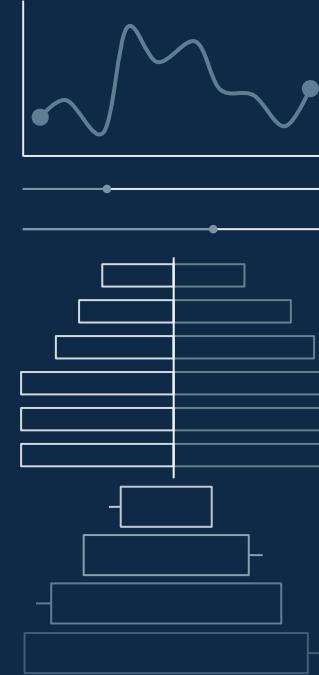
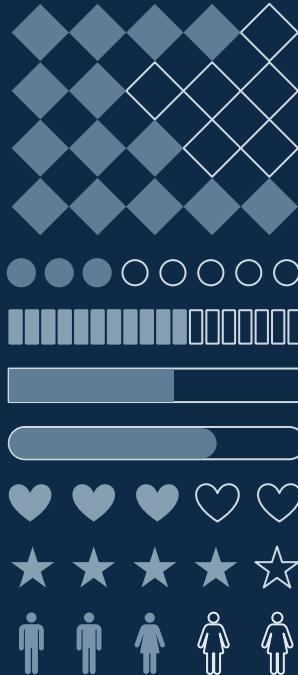
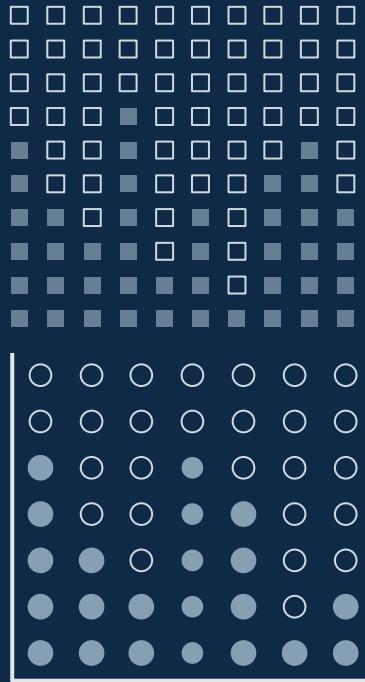












...and our sets of editable icons

You can resize these icons without losing quality.

You can change the stroke and fill color; just select the icon and click on the paint bucket/pen.

In Google Slides, you can also use Flaticon's extension, allowing you to customize and add even more icons.



Educational Icons



Medical Icons



Business Icons



Teamwork Icons



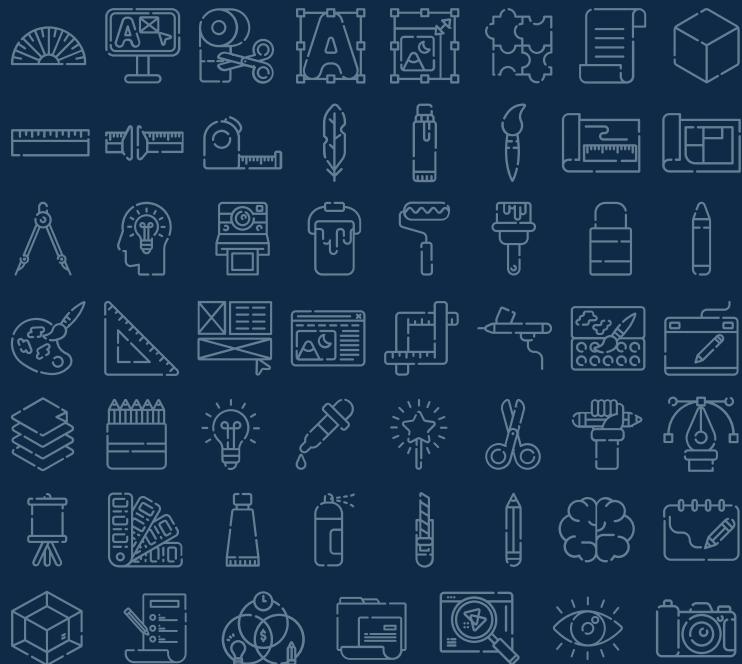
Help & Support Icons



Avatar Icons



Creative Process Icons



Performing Arts Icons



Nature Icons



SEO & Marketing Icons



