

Explorando atenção, impacto, Acesso Aberto e Objetivos de Desenvolvimento Sustentável da investigação a nível mundial. Uma sessão interativa com análise de big data.

A nossa equipa

Mike Taylor

Head of Data Insights



Miguel Garcia

Product Director



Carlos Areia

Senior Data Scientist





Altmetric

Introdução e Metodologia



Dimensions

Part of **DIGITAL**science

Desenvolvimento tecnológico

Parte 1: Introduções e ODS dashboard (9:30 - 10:30)

- Apresentação da equipa
- Apresentação do ODS dashboard
- Exploração do ODS dashboard
- Preparação para a parte 2:
 - Formação equipas parte 2
 - Acesso para a parte 2

Pausa + SETUP (10:30-11:00)

Parte 2 (opcional): Criação visualizações no Looker Studio (11:00 - 13:00)

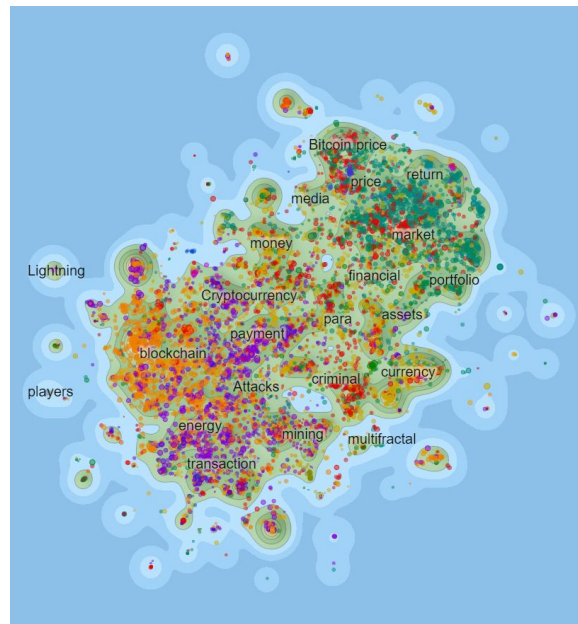
- Apresentação do dataset
- Breve demonstração de como criar visualizações com Looker Studio
- Trabalho em equipa
- Breve discussão e conclusão da sessão

Introdução e Metodologia

Desenvolvimento tecnológico

Crescimento exponencial do poder computacional:

- Big data
- Cloud computing
- Software-as-a-service (SaaS)
- Artificial Intelligence
- Data Science
- Visualisation platforms
- AI
- Etc...

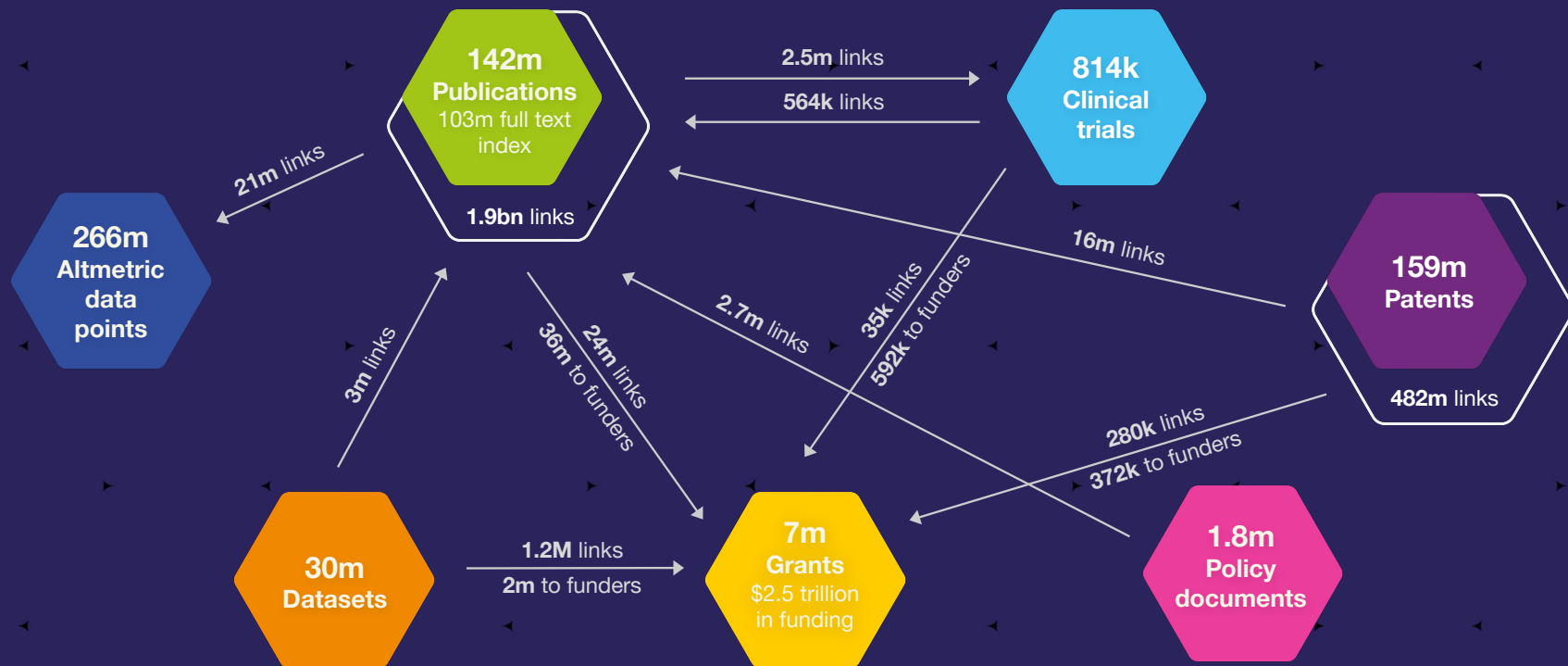


Desenvolvimento tecnológico

O crescimento trouxe novas oportunidades:

- Cálculos que demoravam horas ou dias a processar agora demoram minutos ou segundos
- Projetos e análises que demoravam anos agora demoram horas-minutos.
- Podemos experimentar mais, mais depressa e com maior flexibilidade

We surface the connections



Part of **DIGITAL**science



Sources we track

- Policy documents
- News
- Blogs
- X
- Facebook
- Wikipedia
- Faculty Opinions
- Peer reviews
- Videos
- Patents
- Q&A (Stack overflow)
- Reddit

Historical sources

- Sina Weibo
- Syllabi
- Google+
- LinkedIn
- Pinterest

Pipeline de tratamento de dados



Publicações

Atenção e impacto

Base de dados

Machine learning/AI

Plataforma visual

Parte 1: Plataforma de análise de Objetivos Desenvolvimento Sustentável e Acesso Aberto.

Base de dados

- Informação desde 2015 até ao presente
- 27.600 instituições de educação e saúde
- 7 milhões de publicações, relacionadas ao seu status de Acesso Aberto e respectivos ODS
- 6,5 milhões de investigadores
- 50 milhões de tweets
- 2,5 milhões de notícias
- 1 milhão de postagens no Facebook
- 360.000 postagens de blog
- 240.000 postagens na Wikipedia
- 115.000 patentes
- 55.000 documentos de políticas
- E muito mais



Preparação
de base de
dados



Criação de
dashboard

Base de dados

LatAmerica_publication_dataset

QUERY

SHARE

COPY

SNAPSHOT

DELETE

EXPORT

REFRESH

SCHEMA	DETAILS	PREVIEW	LINEAGE	DATA PROFILE	DATA QUALITY					
Row	pub_id	pub_year	pub_doi	pub_type	pub_date	open_access	open_access_categories_v2	sdg.id	sdg.code	sdg.name
1	pub.1146902638	2022	10.3354/meps14049	article	2022-06-02	No	closed	40014	14	Life Below Water
								40015	15	Life on Land
2	pub.1148095021	2022	10.14201/eks.26763	article	2022-05-23	Yes	oa_all	40004	4	Quality Education
							gold			
3	pub.1092745078	2017	10.4103/pr.pr_124_16	article	2017-01-01	Yes	oa_all	40003	3	Good Health and Well-being
4	pub.1125799601	2020	10.1177/095679762090						5	Gender Equality
5	pub.1135545661	2021	10.1101/2021.02.08.430146	preprint	2021-02-09	Yes	oa_all	40003	3	Good Health and Well-being

of 654153094

Results per page: 50

1 - 50 of 654153094

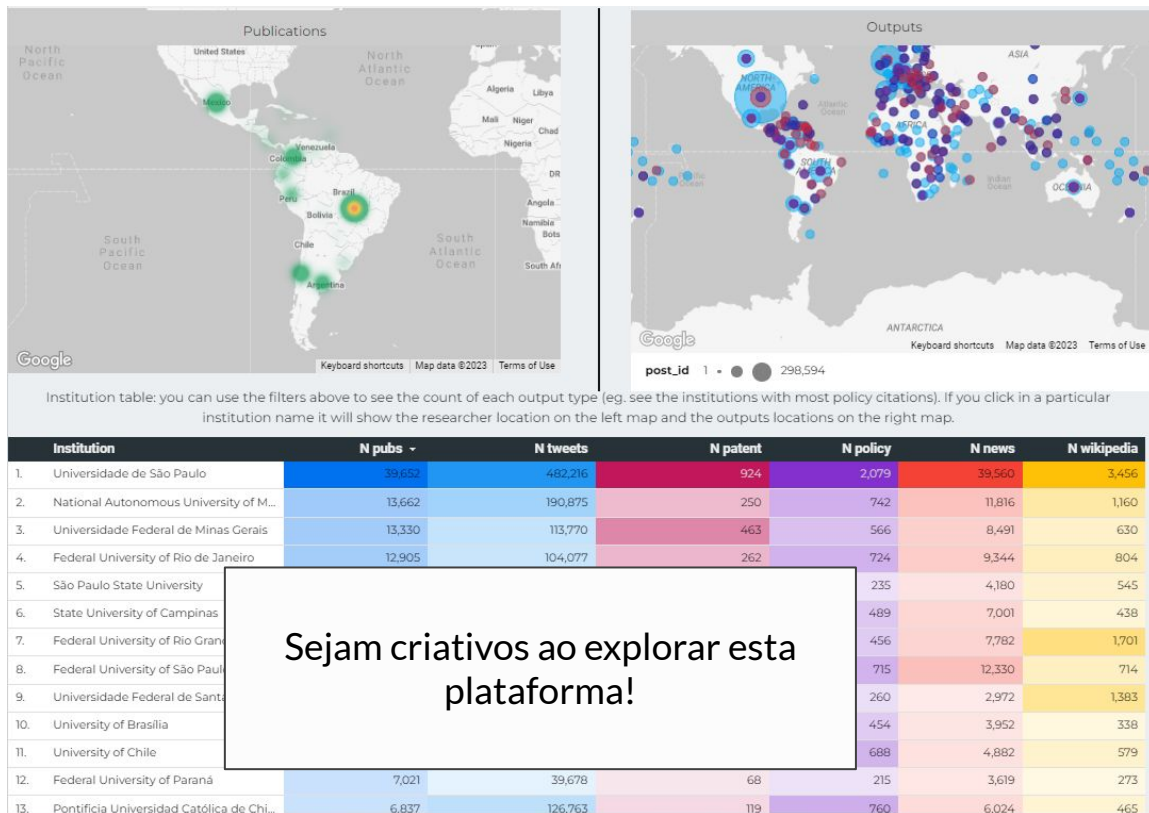
ons

Número de datapoints



Plataforma interativa

- Looker Studio
- O que podem explorar nela?
 - Acesso Aberto
 - ODS
 - Benchmarking
 - Regiões
 - Países
 - Instituições
 - Investigadores
 - Atenção
 - Impacto
 - E muito, mas muito mais!



Preparação para a parte 2

Preparação para a parte 2

Duas opções

Acesso e permissões após o Demo e durante a pausa (~60mins)

Continuar a explorar
o ODS dashboard

Criar o próprio dashboard
com dataset novo (PT-BR)

Equipas entre 3 a 5 pessoas

Como usar a plataforma (DEMO)

Como acceder à plataforma?

<https://shorturl.at/e1n3G>



Disponível 10 dias a partir de hoje!

Demo?

Feedback e contacto

<https://shorturl.at/e1n3G>



- Portátil (perfeito)
- Tablet (ok)
- Smartphone (fraco)



Altmetric

PAUSA + setup



Dimensions

Part of **DIGITAL**science



Altmetric

Parte 2



Dimensions

Part of **DIGITAL**science

Exemplo de uma análise

A liderança da América Latina e Caribe no Acesso Aberto:
uma visão geral da última década.

Objetivos

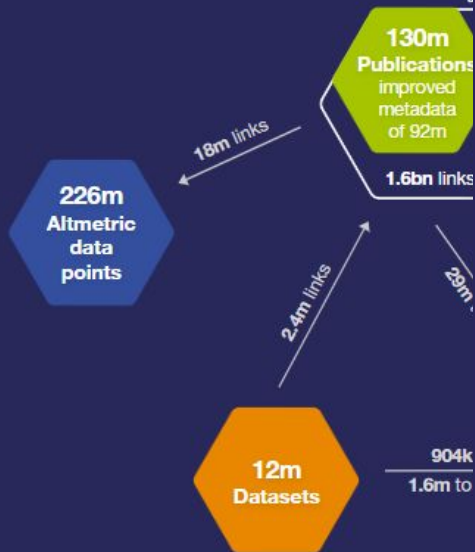
Fazer uma análise aprofundada das tendências no Acesso Aberto:

- Desenvolvimento da America Latina na última década
- Comparar com o resto do mundo
- Explorar a desenvolvimento e contribuição de cada país
- Explorar também a nível institucional
- Etc...

Metodologia

Dimensions

Altmetric - sources

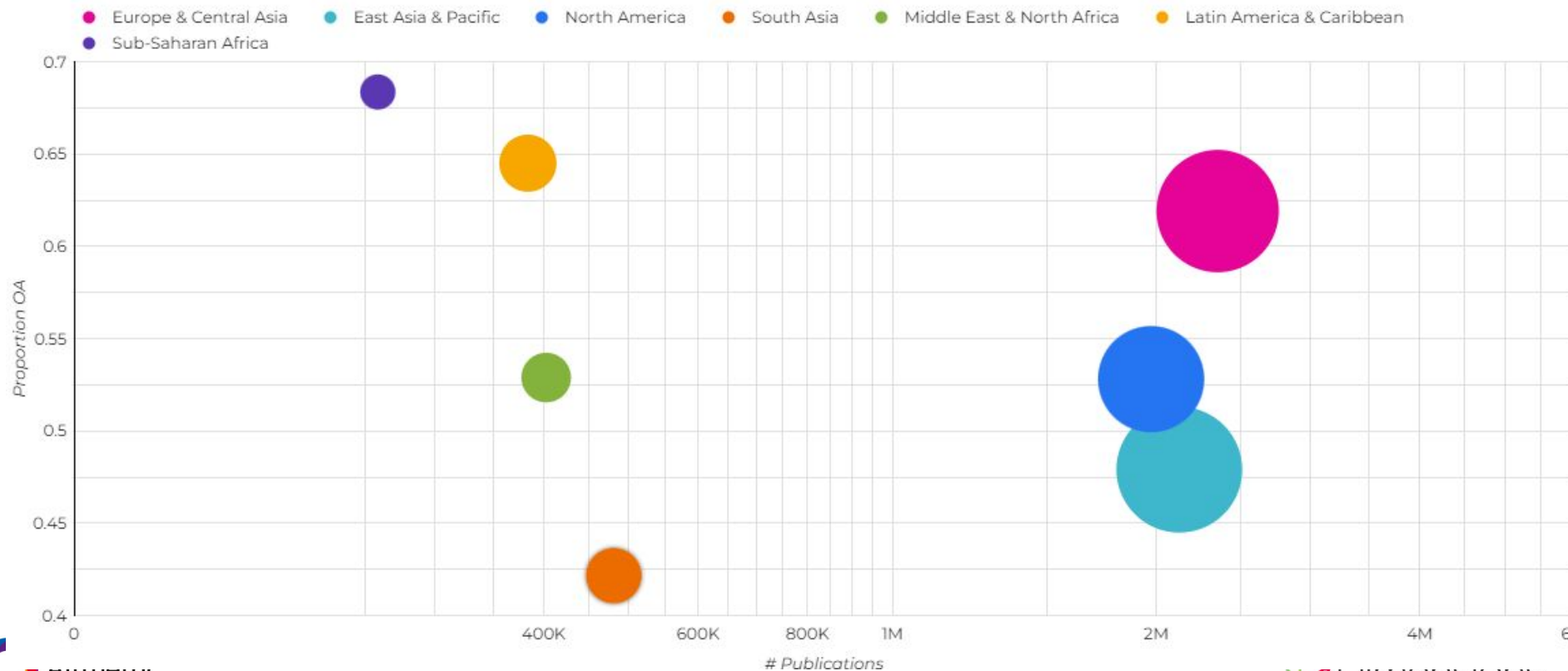


- Policy documents
- News
- Blogs
- Twitter
- Facebook
- Wikipedia
- Faculty Opinions
- Peer Reviews
- Videos
- Patents
- Q&A (Stack Overflow)
- Reddit

Historical sources

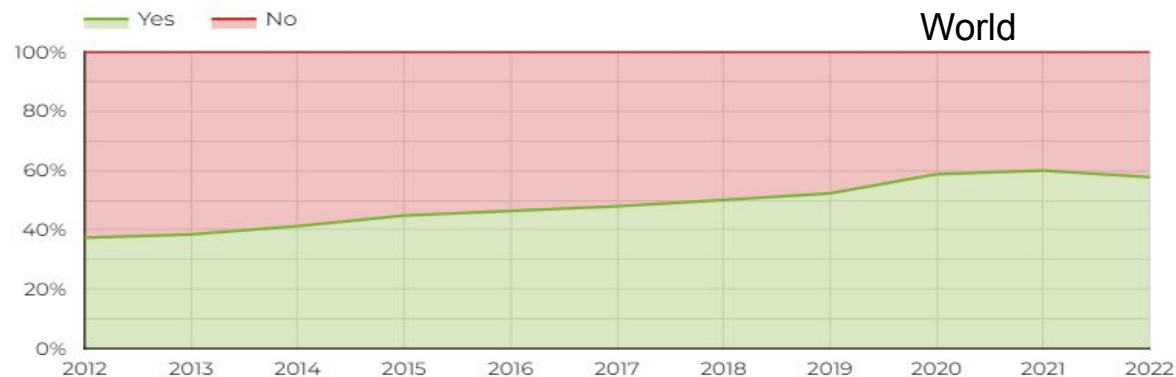
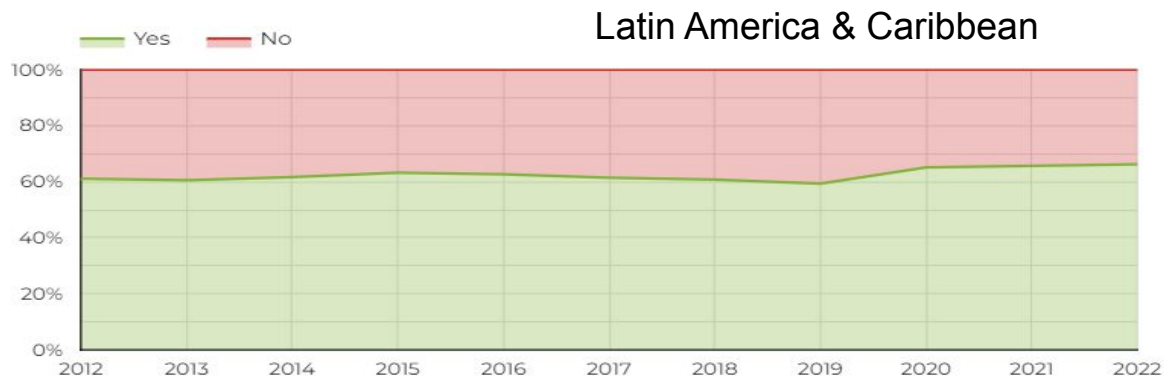
- Sina Weibo
- Syllabi
- Google+
- LinkedIn
- Pinterest

Publicações - proporção AA última década

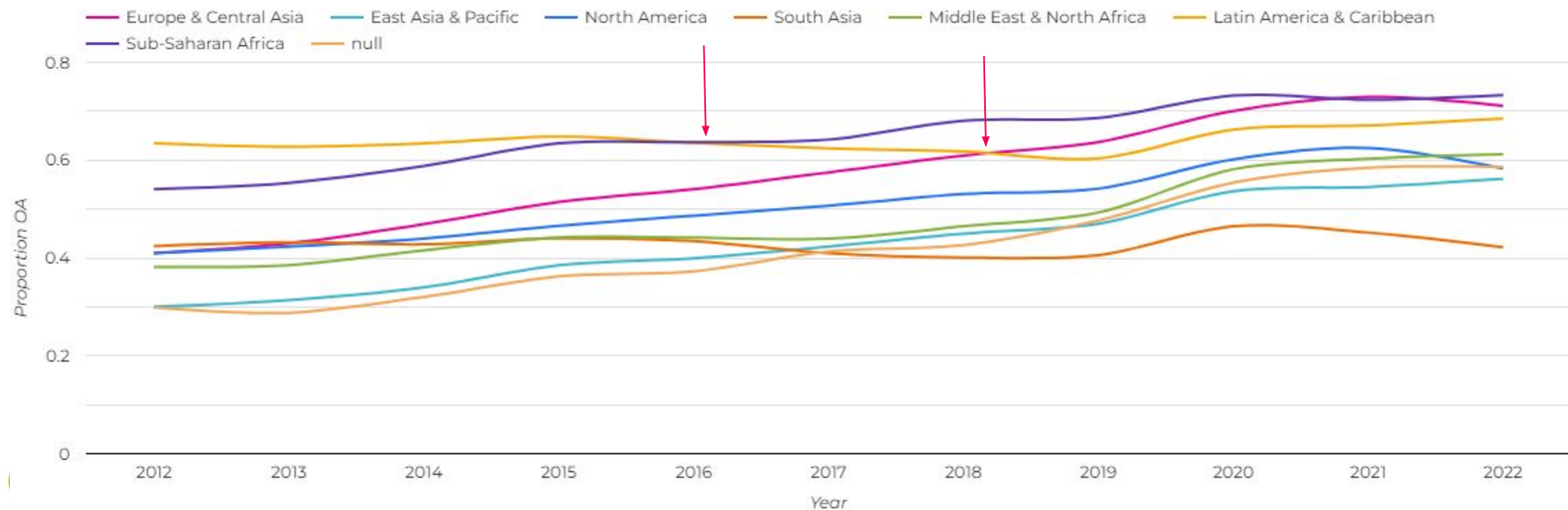


Desenvolvimento AA - panorama geral

- Proporção AA na ALC maior que a média mundial em todos os anos.
- Especialmente no início (2012)

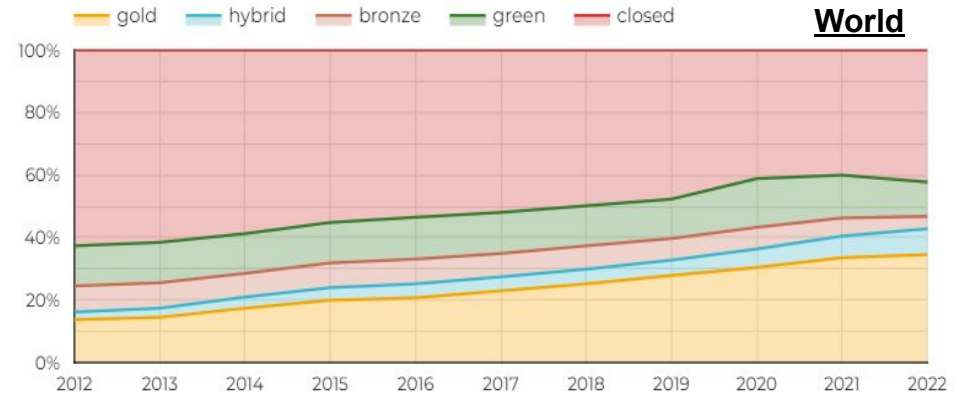
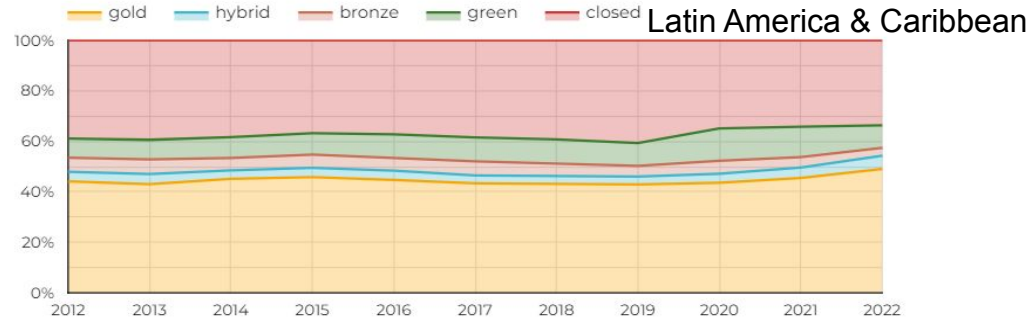


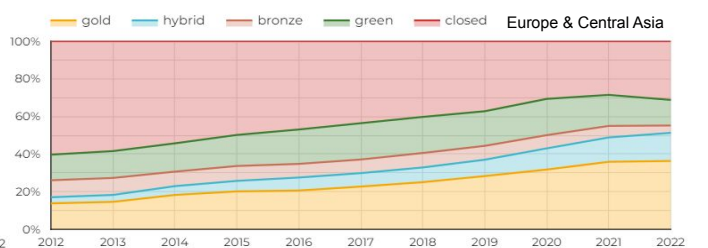
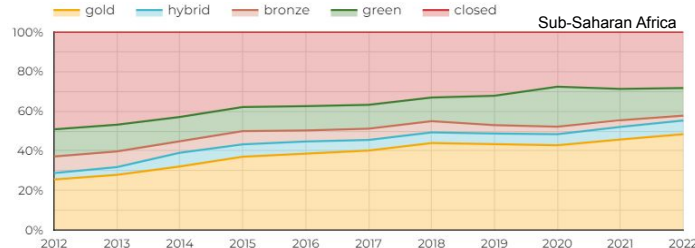
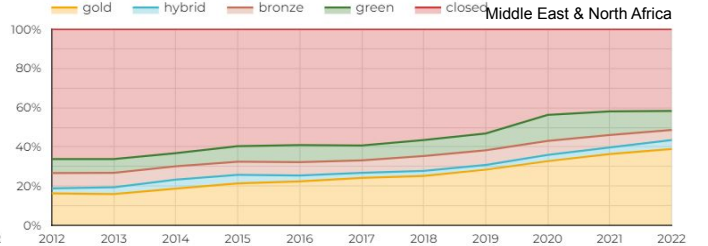
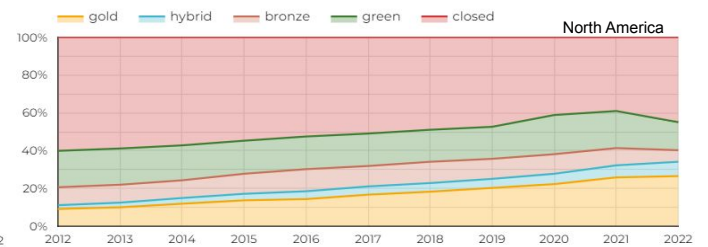
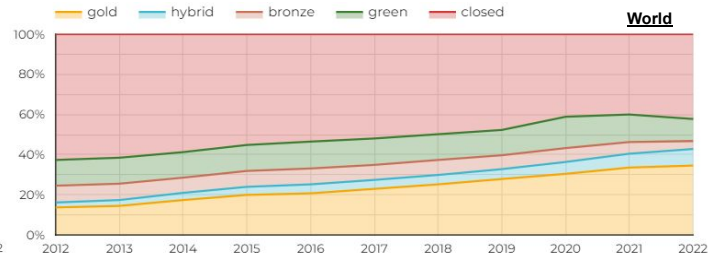
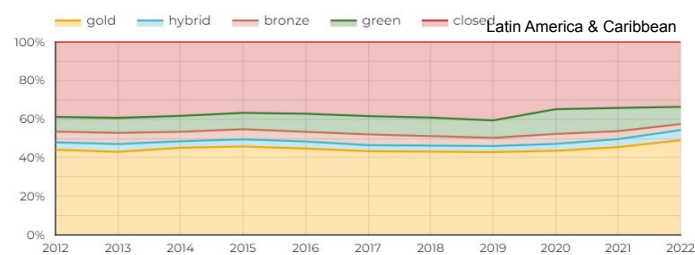
Desenvolvimento AA - panorama geral



Desenvolvimento AA - parte 2

- Também no gold AA a América Latina se destaca desde o início (44%).
- Nota-se um crescimento mais acentuado nas outras regiões
- Aproximando-se da América Latina nos últimos anos.

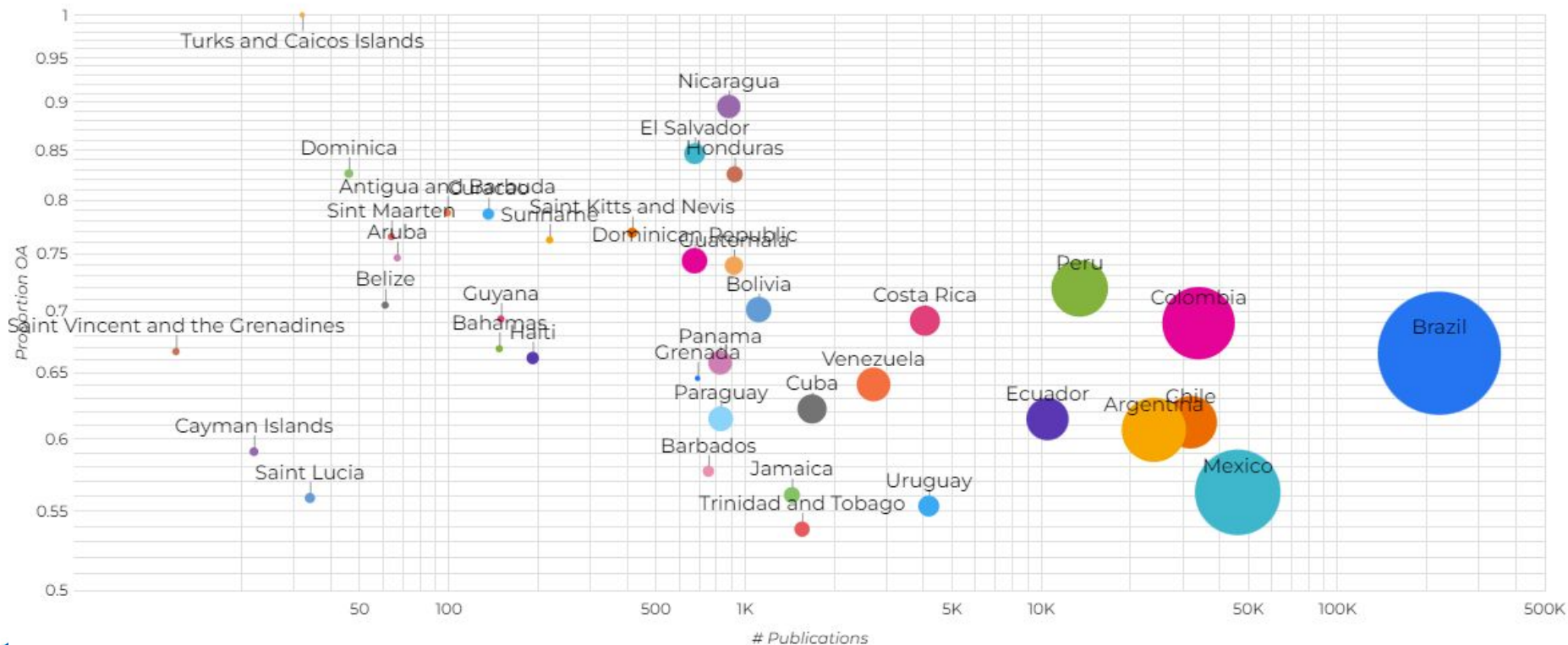




Resultados

Dentro da América Latina

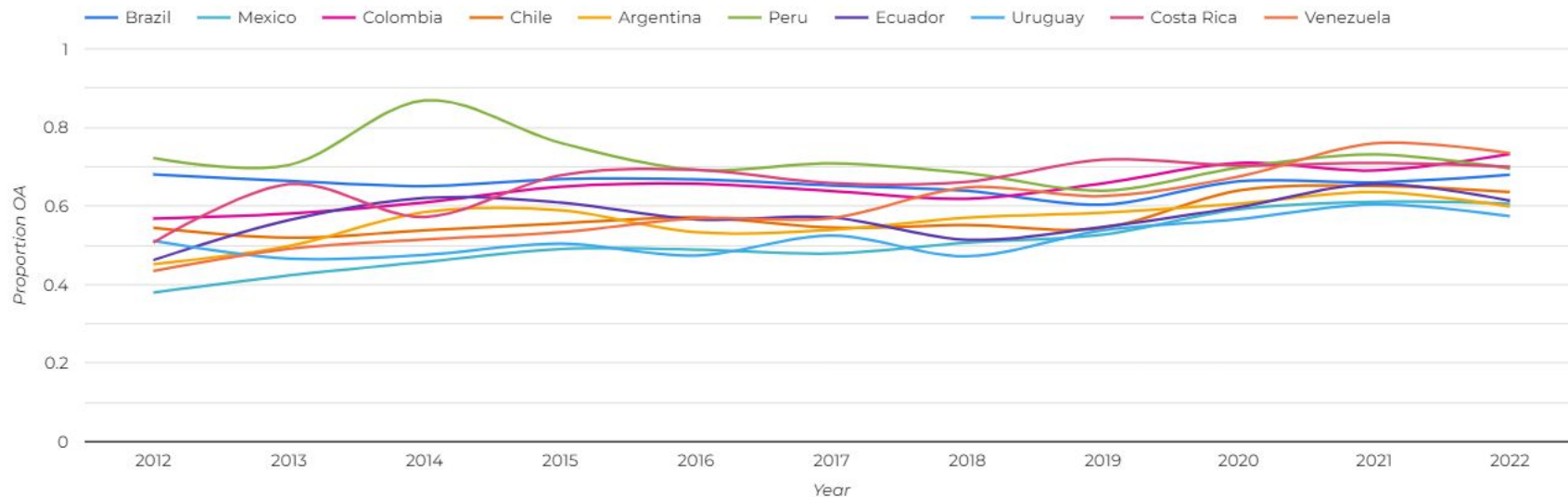
AA e volume publicações por país - ALC



Proporção AA por país na última década

Year / Proportion OA (n pubs)											
Country	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Antigua and Barbuda	-	-	100% (1)	-	100% (1)	88.89% (9)	60% (5)	100% (1)	83.33% (6)	77.27% (22)	75% (24)
Argentina	55.84% (1001)	61.36% (1131)	67.96% (1214)	66.05% (1355)	54.66% (1480)	55.07% (1816)	58.42% (1936)	59.26% (2187)	61.73% (2971)	64.68% (3273)	62.81% (3251)
Aruba	-	100% (2)	75% (4)	25% (4)	66.67% (6)	100% (2)	100% (4)	80% (5)	80% (15)	50% (4)	81.82% (11)
Bahamas	0% (1)	80% (5)	90.91% (11)	55.56% (9)	83.33% (6)	50% (8)	62.5% (8)	71.43% (14)	76.92% (13)	75.76% (33)	56.52% (23)
Barbados	22.58% (31)	50% (28)	41.67% (36)	53.85% (52)	48.84% (43)	52.63% (38)	62.5% (64)	60.56% (71)	70.27% (74)	59.63% (109)	67.59% (108)
Belize	-	-	-	75% (4)	42.86% (7)	66.67% (3)	25% (8)	100% (6)	100% (5)	100% (7)	75% (12)
Bolivia	56.52% (23)	61.82% (55)	59.26% (54)	63.77% (69)	52.94% (68)	57.97% (69)	69.05% (84)	78.35% (97)	69.57% (138)	74.84% (155)	80.81% (172)
Brazil	69.33% (9331)	67.4% (10358)	65.96% (11558)	67.61% (13107)	67.42% (15143)	65.98% (16365)	64.72% (18400)	61.15% (19272)	67.61% (26251)	67.49% (30116)	69.8% (28816)
Cayman Islands	0% (1)	100% (1)	-	-	-	0% (2)	100% (3)	0% (1)	33.33% (3)	60% (5)	100% (1)
Chile	56.05% (992)	53.76% (1198)	55.29% (1409)	58.23% (1743)	58.17% (2032)	55.7% (2201)	56.34% (2460)	55.92% (2874)	64.88% (3983)	66.55% (4762)	66.54% (4593)
Colombia	63.02% (1033)	62.86% (1131)	64.5% (1372)	67.46% (1589)	66.68% (1918)	64.65% (2433)	63.24% (2783)	66.68% (3349)	71.86% (4502)	70.92% (5123)	75.7% (4913)
Costa Rica	51.33% (150)	68.51% (181)	57.71% (201)	67.38% (233)	70.54% (241)	65.56% (270)	65.84% (322)	73.02% (367)	72.14% (524)	71.4% (584)	71.75% (531)
Cuba	45.71% (70)	41.18% (85)	50% (86)	54.93% (71)	51.96% (102)	55.45% (110)	60% (135)	66.42% (137)	65.98% (194)	73.39% (248)	69.65% (257)
Curacao	50% (4)	57.14% (7)	83.33% (6)	77.78% (9)	100% (5)	100% (9)	77.78% (9)	54.55% (11)	83.33% (12)	95.24% (21)	77.27% (22)
Dominica	-	-	-	-	0% (1)	-	-	100% (4)	70% (10)	81.82% (11)	85.71% (14)
Dominican Republic	70.59% (17)	62.5% (8)	50% (18)	76.19% (21)	66.67% (24)	82.14% (28)	76.32% (38)	71.43% (63)	72.28% (101)	79.85% (134)	71.88% (128)
Ecuador	51.26% (119)	57.95% (176)	62.31% (260)	61.52% (330)	57.59% (507)	58.35% (701)	52.83% (884)	56.61% (1014)	60.95% (1470)	66.2% (1716)	65.54% (1744)
El Salvador	93.75% (16)	64.29% (14)	67.86% (28)	77.78% (18)	96.88% (32)	90.28% (72)	75% (48)	87.69% (65)	79.55% (88)	84.93% (73)	86.81% (144)
Grenada	33.33% (15)	25% (20)	40.74% (27)	63.33% (30)	46.51% (43)	58.06% (62)	71.93% (57)	58.49% (53)	69.66% (89)	71.43% (105)	81.42% (113)

Proporção AA por país na última década





Altmetric

0 dataset



Dimensions

Part of **DIGITAL**science

The dataset

- Apenas Brasil e Portugal
 - >250k publicações
 - 670k profiles
 - ~2m tweets
 - ~150k noticias
 - 30k facebook
 - 10k wikipedia
 - 5k politicas
 - 3k patents
 - Etc...
-

The dataset

- **pub_id** - Dimensions publication id (eg. pub_76374858)
 - **pub_title**
 - **pub_year**
 - **pub_type** - (eg. article, preprint, etc...)
 - **open_access** (Yes = open, No = closed)
 - **SDG** (ODS)
 - **organisation_id** - Dimensions institution id
 - **organisation_name**
 - **organisation_country**
 - **organisation_type**
 - **altmetric_id**
 - **altmetric_score**
 - **citations_count**
 - **post_id** (Altmetric mention ID)
 - **post_year**
 - **post_country**
 - **post_type** (eg. tweet, wikipedia, policy, etc...)
 - **post_profile_id**
-



Altmetric

Looker Studio



Dimensions

Part of **DIGITAL**science

- Um pré requisito é saber as bases do Looker Studio
- Aqui está um start guide se for necessário durante o workshop

https://support.google.com/looker-studio/answer/9171315?hl=en&ref_topic=6267740&sjid=5547410759238179458-EU#zippy=



Tip 1: Quando múltiplas pessoas estão a trabalhar no mesmo report trabalhem em páginas diferentes

Tip 2: Se o dashboard estiver muito lento, criem um filtro e pré-seleccionem algo para diminuir o tamanho do dataset (ex. Filtrar um país ou afiliação)

Tip 3: Usem o CTD (count distinct) para evitar counts duplicados

Conselhos

confOA_dataset								
<div>QUERY SHARE COPY SNAPSHOT DELETE EXPORT REFRESH</div>								
SCHEMA DETAILS PREVIEW TABLE EXPLORER PREVIEW INSIGHTS LINEAGE DATA PROFILE DATA QUALITY								
Row	pub_id	pub_year	pub_type	open_access	sdg	pub_title to pollution	organis	
3	pub.1000020868	2017	article	No	12 Responsible Consumption a...	A design tool to diagnose product recyclability during product design phase	grid.411	
4	pub.1000020868	2017	article	No	12 Responsible Consumption a...	A design tool to diagnose product recyclability during product design phase	grid.412	
5	pub.1000023905	2016	article	Yes	3 Good Health and Well Being	Global Health Policy and Access to Care: Investigating Patient Choice on an International Level Using Social Media	grid.839	
6	pub.1000023905	2016	article	Yes	3 Good Health and Well Being	Global Health Policy and Access to Care: Investigating Patient Choice on an International Level Using Social Media	grid.839	
7	pub.1000023905	2016	article	Yes	3 Good Health and Well Being	Global Health Policy and Access to Care: Investigating Patient Choice on an International Level Using Social Media	grid.839	

- **Objetivos:**
 - Escolher um nome para a vossa equipa
 - Escolher um nome e objetivo para o vosso dashboard
 - Serem creativos e divertirem-se
 - Todos os dashboards (ODS e exemplos criados) estarão disponíveis e acessíveis até ao dia 14 de Outubro (apenas para visualização, não para edição)
 - Até esta data podem partilhar os dashboards também com quem quiserem.
-

**Exemplo como criar
visualizações no
LookerStudio?**



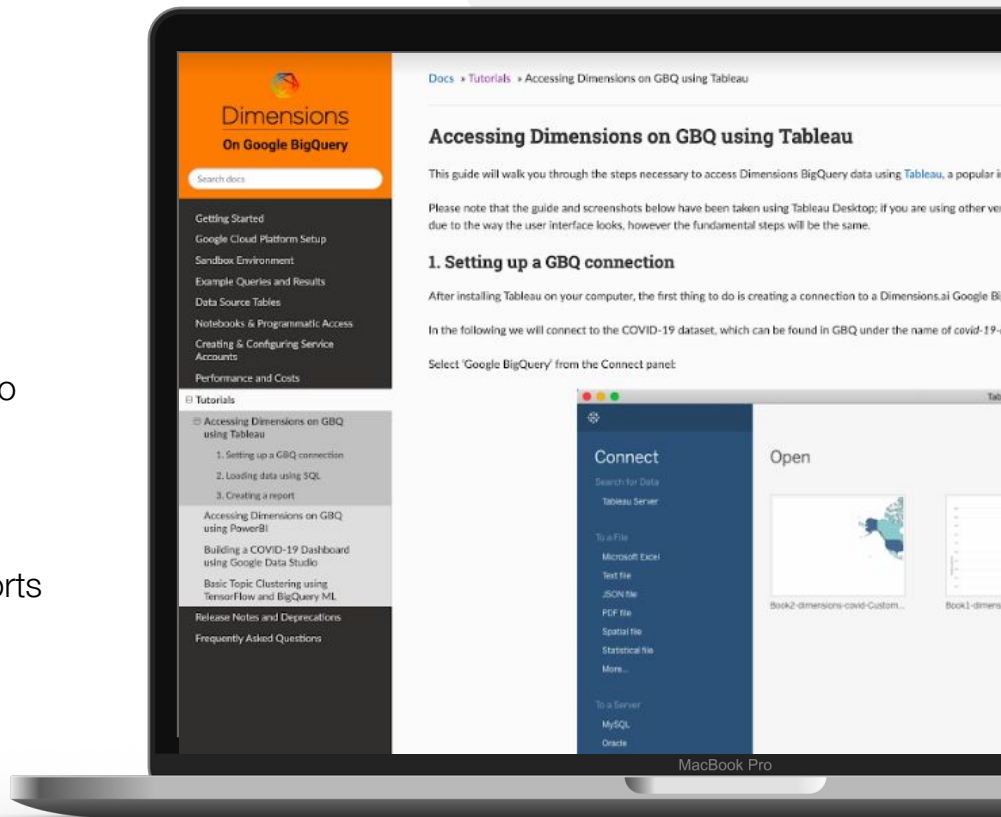
Têm até às 12:45!!!

Discussão/Conclusão

Dimensions on Google BigQuery

Access to the full Dimensions database,
via a direct integration with BigQuery -
Google's cloud data warehouse

- **Integrate** Dimensions data into your existing reporting and analysis infrastructure quickly
- **Analyze** the full dataset with complete flexibility to support multiple decision-making processes
- **Join** Dimensions data to your own private data
- **Create** custom dashboards and automated reports on different topics
- **Share** these analyses and dashboards with stakeholders across the organization





Altmetric on Google BigQuery

A fuller set of data for deeper, richer insights

Integrate Altmetric data with internal datasets



Combine Altmetric data with other available datasets



Pull Altmetric data into internal dashboards and business intelligence tools



- Run complex, enriched queries using the depth of Altmetric data via Google BigQuery
- Explore unseen connections and create the actionable insights you need.
- Analyze billions of data points in seconds
- See the full picture of your organization's influence and performance



@C_Areia



carlos@altmetric.com

Dashboards disponíveis até 14 Outubro.
Podem partilhar o link até essa data