



**GET YOUR GUIDE**





**How far is the Empire State Building  
from the Eiffel Tower?**

**0.22**

**Haystack 2023-04-25**

**Ansgar Gruene**

**GET YOUR GUIDE**



# Introduction



GET YOUR GUIDE

# Author



## Ansgar Gruene

Senior Data Scientist,  
Traveler Data Products team

[ansgar.gruene@getyourguide.com](mailto:ansgar.gruene@getyourguide.com)



A person with long dark hair, wearing a bright red jacket and dark pants, stands in the foreground looking up at a large, multi-tiered waterfall. The waterfall is the central focus, with water cascading down dark, mossy rocks. The surrounding forest is dense with green foliage and some autumn-colored leaves. The scene is captured with a long exposure, giving the water a soft, ethereal appearance. The overall mood is serene and majestic.

# GET YOUR GUIDE in a nutshell



Chase the Northern Lights in Norway  
around a fire with a hot drink



Experience a tour of the vibrant MaeKlong  
Railway Market and Damnoen Saduak  
Floating Market in Bangkok



## We unlock the world's most unforgettable travel experiences

Go on a desert safari in Dubai  
with sandboarding and  
a camel ride



GET  
YOUR  
GUIDE

Enjoy a delicious experience at the  
Museum of Ice Cream in New York





# We're the **leading global marketplace** for travel experiences



**75k+**

curated tours  
& activities



**80m+**

tickets sold



**\$883m+**

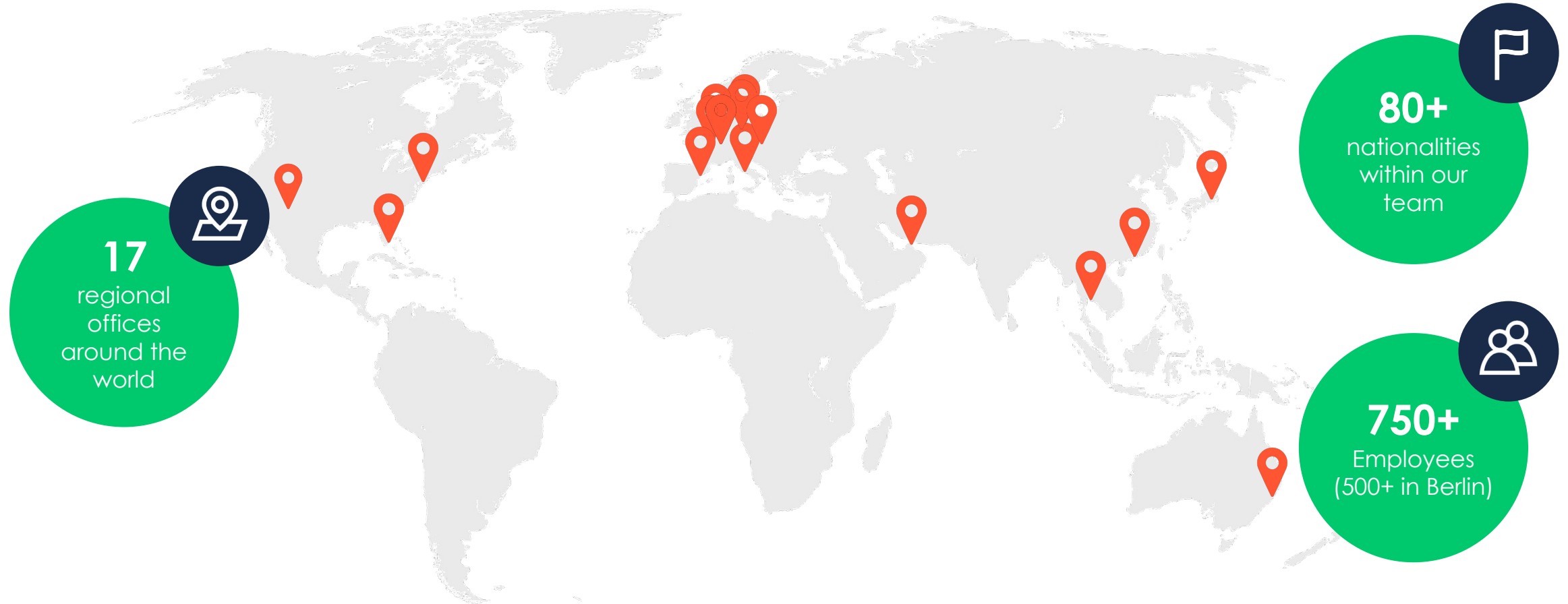
in total  
funding



**16k+**

suppliers

# Unforgettable experiences come from all over the world, **so do we.**





Unforgettable experiences come from all over the world, **so do we.**



# Unforgettable experiences come from all over the world, **so do we.**

A world map with several red location pins. Two blue arrows point from a cluster of pins in Europe to two inset photographs. The first inset shows the Brandenburg Gate in Berlin. The second inset shows a group of men in white shirts and suspenders, holding large beer mugs and raising them in a toast at a beer festival.

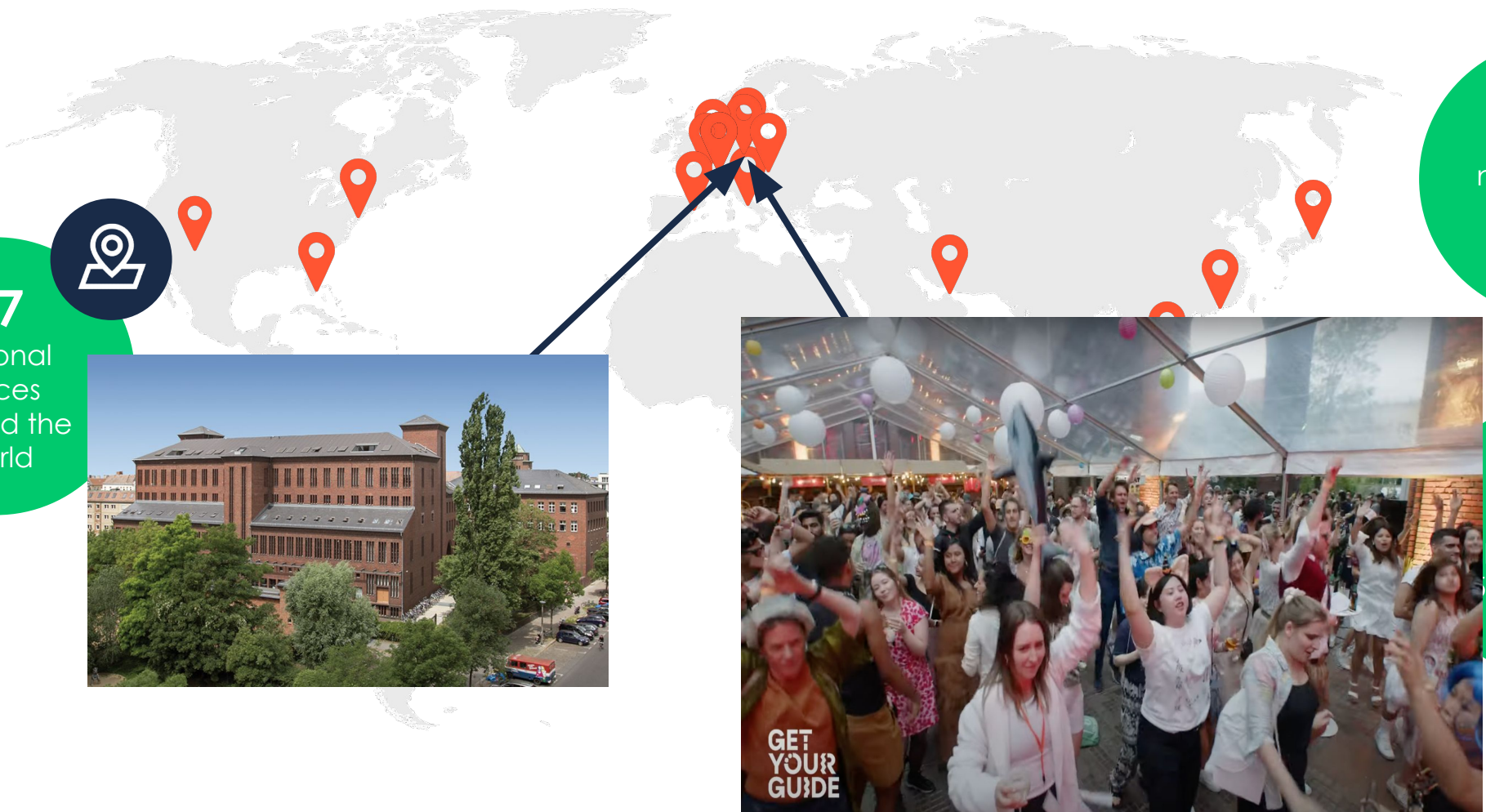
**17**  
regional  
offices  
around the  
world

**80+**  
nationalities  
within our  
team

**750+**  
Employees  
(100+ in Berlin)



# Unforgettable experiences come from all over the world, so do we.



A world map with 17 red location pins. Two blue arrows point from the Berlin area to two inset photos. The left photo shows a large brick building with a tower. The right photo shows a large group of people at a party with balloons and a shark mascot. The text 'GET YOUR GUIDE' is visible in the bottom left of the party photo.

**17** regional offices around the world

**80+** nationalities within our team

**750+** Employees (100+ in Berlin)

Unforgettable experiences come from all over the world, **so do we.**







# GetYourGuide + Semantic Vectors



New York City

# Viewing points

Price ▾

Languages ▾

Duration ▾

Time ▾

Filters

29 activities found ⓘ

Sort by: Recommended ▾



ENTRY TICKET  
NYC: SUMMIT One Vanderbilt Experience Ticket

404501

2 hours

Top pick

Booked 170 times yesterday

★★★★★ 4.8 (5,989 reviews)

From € 41.68 per person



ENTRY TICKET  
NYC: Empire State Building Tickets & Skip-the-Line

6195

45 minutes - 1 hour • Optional audio guide

Booked 62 times yesterday

★★★★★ 4.6 (8,009 reviews)

From € 43.66 per person



ENTRY TICKET  
NYC: Edge Observation Deck Admission Ticket

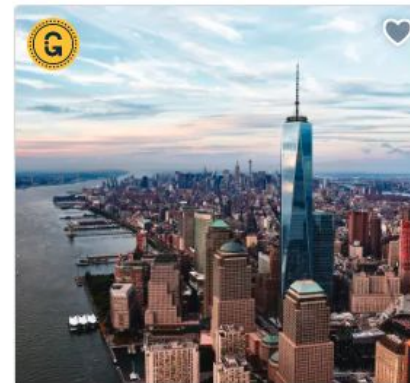
404016

1 day

Booked 43 times yesterday

★★★★★ 4.6 (2,217 reviews)

From € 41.67 per person



ENTRY TICKET  
New York One World Observatory: Skip-the-Line Ticket Options

52132

1 day • Skip the line

★★★★★ 4.7 (7,917 reviews)

From € 43.66 per person



Page 1 of 2



GET YOUR GUIDE



New York City

# Viewing points

Price ▾

Languages ▾

Duration ▾

Time ▾

Filters

29 activities found ⓘ

Sort by: Recommended ▾



ENTRY TICKET  
NYC: SUMMIT One Vanderbilt Experience Ticket

404501

2 hours

Top pick

Booked 170 times yesterday

★★★★★ 4.8 (5,989 reviews)

From € 41.68 per person



ENTRY TICKET  
NYC: Empire State Building Tickets & Skip-the-Line

6195

45 minutes - 1 hour • Optional audio guide

Booked 62 times yesterday

★★★★★ 4.6 (8,009 reviews)

From € 43.66 per person



ENTRY TICKET  
NYC: Edge Observation Deck Admission Ticket

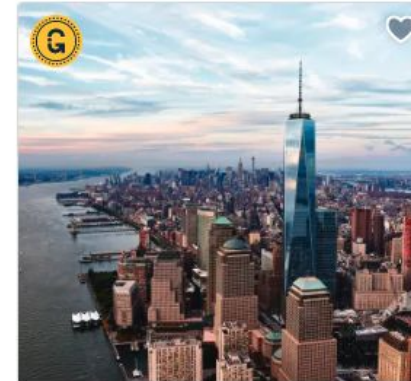
404016

1 day

Booked 43 times yesterday

★★★★★ 4.6 (2,217 reviews)

From € 41.67 per person



ENTRY TICKET  
New York One World Observatory: Skip-the-Line Ticket Options

52132

1 day • Skip the line

★★★★★ 4.7 (7,917 reviews)

From € 43.66 per person



Page 1 of 2



GET YOUR GUIDE

New York City

# Viewing points

Price ▾

Languages ▾

Duration ▾

Time ▾

Filters

29 activities found ⓘ

Sort by: Recommended ▾



ENTRY TICKET  
NYC: SUMMIT One Vanderbilt Experience Ticket

404501

2 hours

Top pick

Booked 170 times yesterday

★★★★★ 4.8 (5,989 reviews)

From € 41.68 per person



ENTRY TICKET  
NYC: Empire State Building Tickets & Skip-the-Line

6195

45 minutes - 1 hour • Optional audio guide

Booked 62 times yesterday

★★★★★ 4.6 (8,009 reviews)

From € 43.66 per person



ENTRY TICKET  
NYC: Edge Observation Deck Admission Ticket

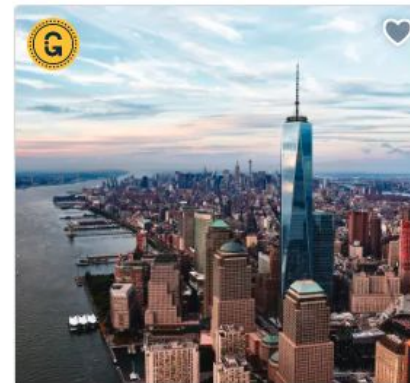
404016

1 day

Booked 43 times yesterday

★★★★★ 4.6 (2,217 reviews)

From € 41.67 per person



ENTRY TICKET  
New York One World Observatory: Skip-the-Line Ticket Options

52132

1 day • Skip the line

★★★★★ 4.7 (7,917 reviews)

From € 43.66 per person



Page 1 of 2



GET YOUR GUIDE



# New York City Viewing points

Price ▾ Languages ▾ Duration ▾ Time ▾ Filters

29 activities found ⓘ

Sort by: Recommended ▾



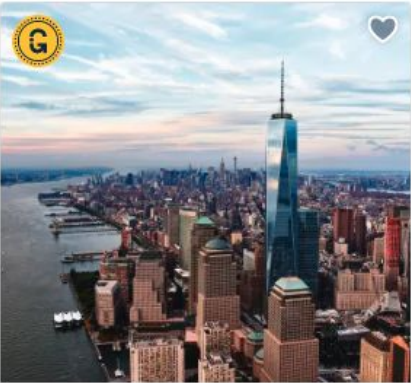
**ENTRY TICKET**  
**NYC: SUMMIT One Vanderbilt Experience Ticket**  
404501  
2 hours  
Top pick  
Booked 170 times yesterday  
★★★★★ 4.8 (5,989 reviews)  
From € 41.68 per person



**ENTRY TICKET**  
**NYC: Empire State Building Tickets & Skip-the-Line**  
6195  
45 minutes - 1 hour • Optional audio guide  
Booked 62 times yesterday  
★★★★★ 4.6 (8,009 reviews)  
From € 43.66 per person



**ENTRY TICKET**  
**NYC: Edge Observation Deck Admission Ticket**  
404016  
1 day  
Booked 43 times yesterday  
★★★★★ 4.6 (2,217 reviews)  
From € 41.67 per person



**ENTRY TICKET**  
**New York One World Observatory: Skip-the-Line Ticket Options**  
52132  
1 day • Skip the line  
★★★★★ 4.7 (7,917 reviews)  
From € 43.66 per person

# Things to do in Paris

A city so synonymous with beauty, even the least romantic are defenseless against its charms.



## Available activities

Price ▾

Languages ▾

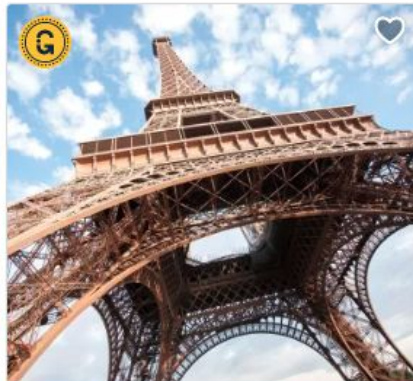
Duration ▾

Time ▾

Filters

797 activities found ⓘ

Sort by: Recommended ▾



**GUIDED TOUR**  
Paris: Eiffel Tower Summit or Second Floor Access

403969

Top pick

★★★★★ 4.3 (12,178 reviews)

From € 40

From € 30 per person



**ENTRY TICKET**  
Disneyland Paris 1-Day Ticket

395320

★★★★★ 4.5 (16,673 reviews)

From € 56 per person



**PRIVATE TOUR**  
Private 5-hour tour to Versailles from Paris

447845

New activity

From € 597 per person



**HOP-ON HOP-OFF TOUR**  
Paris: Tootbus Hop-on Hop-off Discovery Bus Tour

381474

★★★★★ 4.2 (2,582 reviews)

From € 42

From € 35.70 per person



# Things to do in Paris

A city so synonymous with beauty, even the least romantic are defenseless against its charms.



## Available activities

Price ▾

Languages ▾

Duration ▾

Time ▾

Filters

797 activities found ⓘ

Sort by: Recommended ▾




**GUIDED TOUR**  
Paris: Eiffel Tower Summit or Second Floor Access

403969

Top pick

★★★★★ 4.3 (12,178 reviews)

From € 40  
From € 30 per person




**ENTRY TICKET**  
Disneyland Paris 1-Day Ticket

395320

★★★★★ 4.5 (16,673 reviews)

From € 56 per person



**PRIVATE TOUR**  
Private 5-hour tour to Versailles from Paris

447845

New activity

From € 597 per person



**HOP-ON HOP-OFF TOUR**  
Paris: Tootbus Hop-on Hop-off Discovery Bus Tour

381474

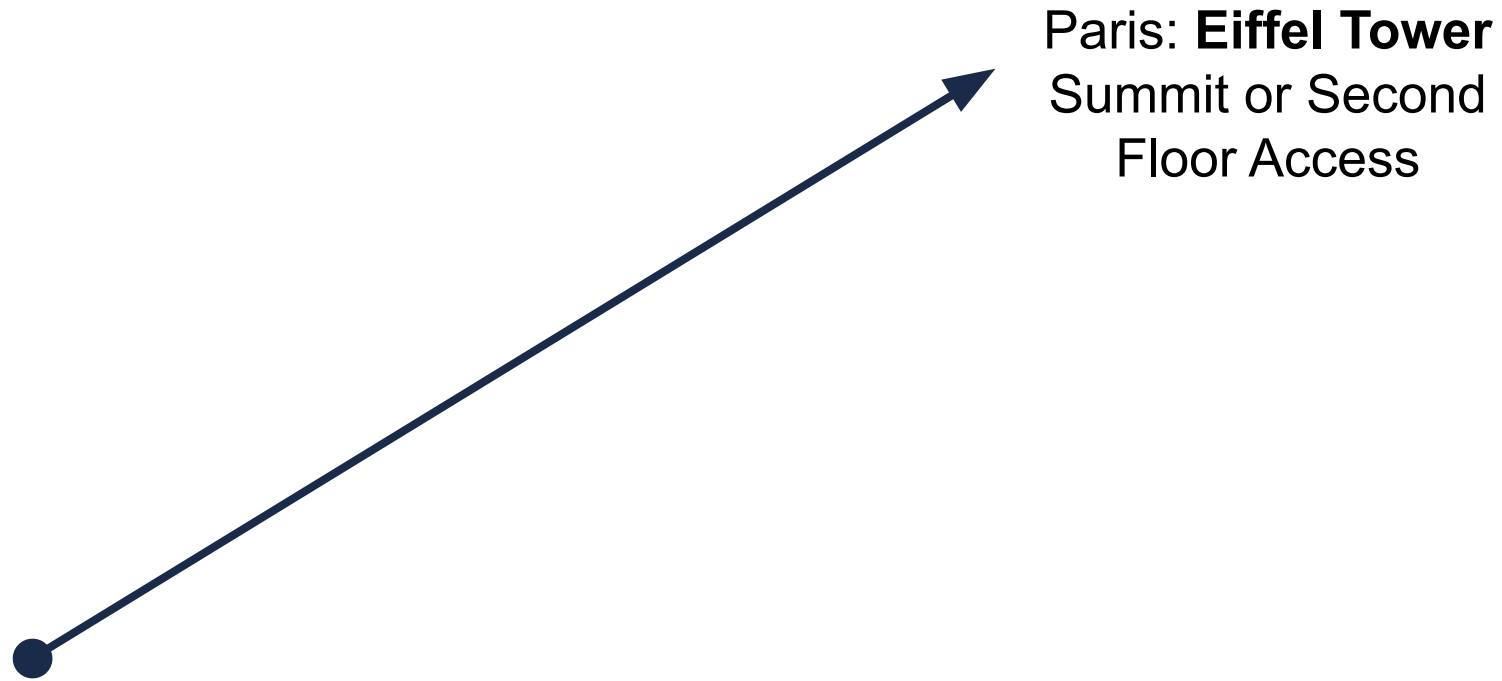
★★★★★ 4.2 (2,582 reviews)

From € 42  
From € 35.70 per person

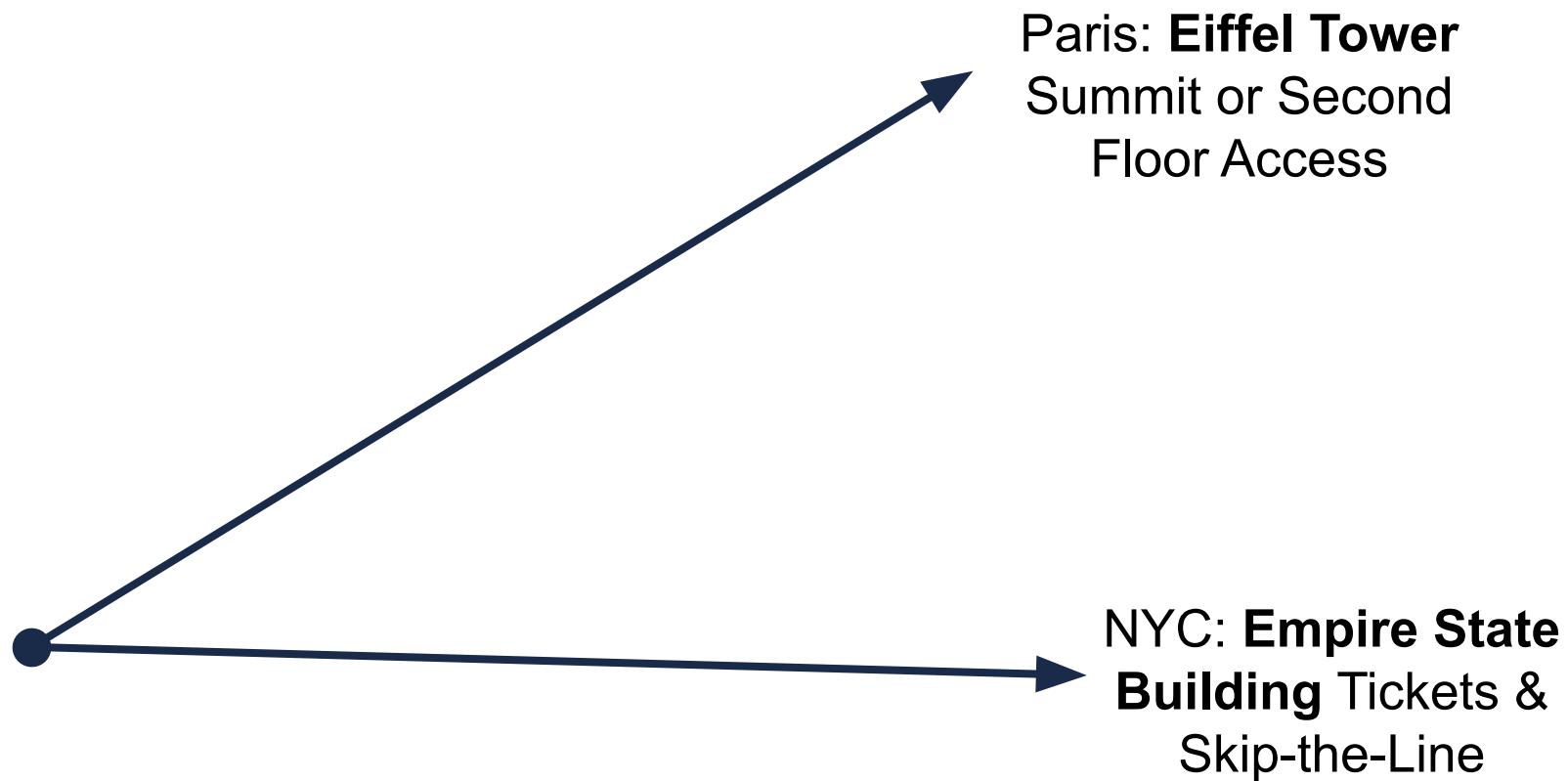
# Semantic Vectors at GetYourGuide: Example



# Semantic Vectors at GetYourGuide: Example

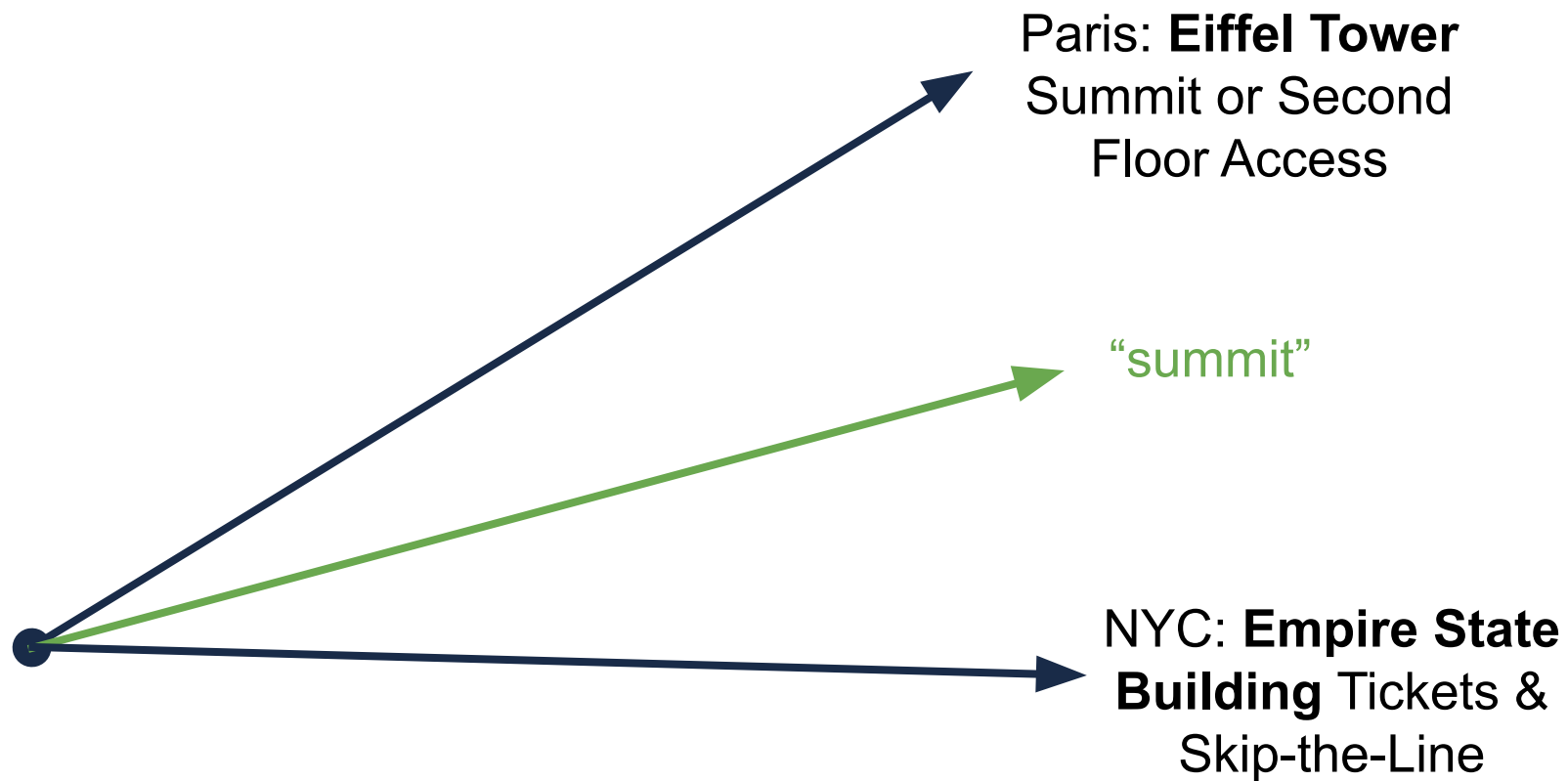


# Semantic Vectors at GetYourGuide: Example

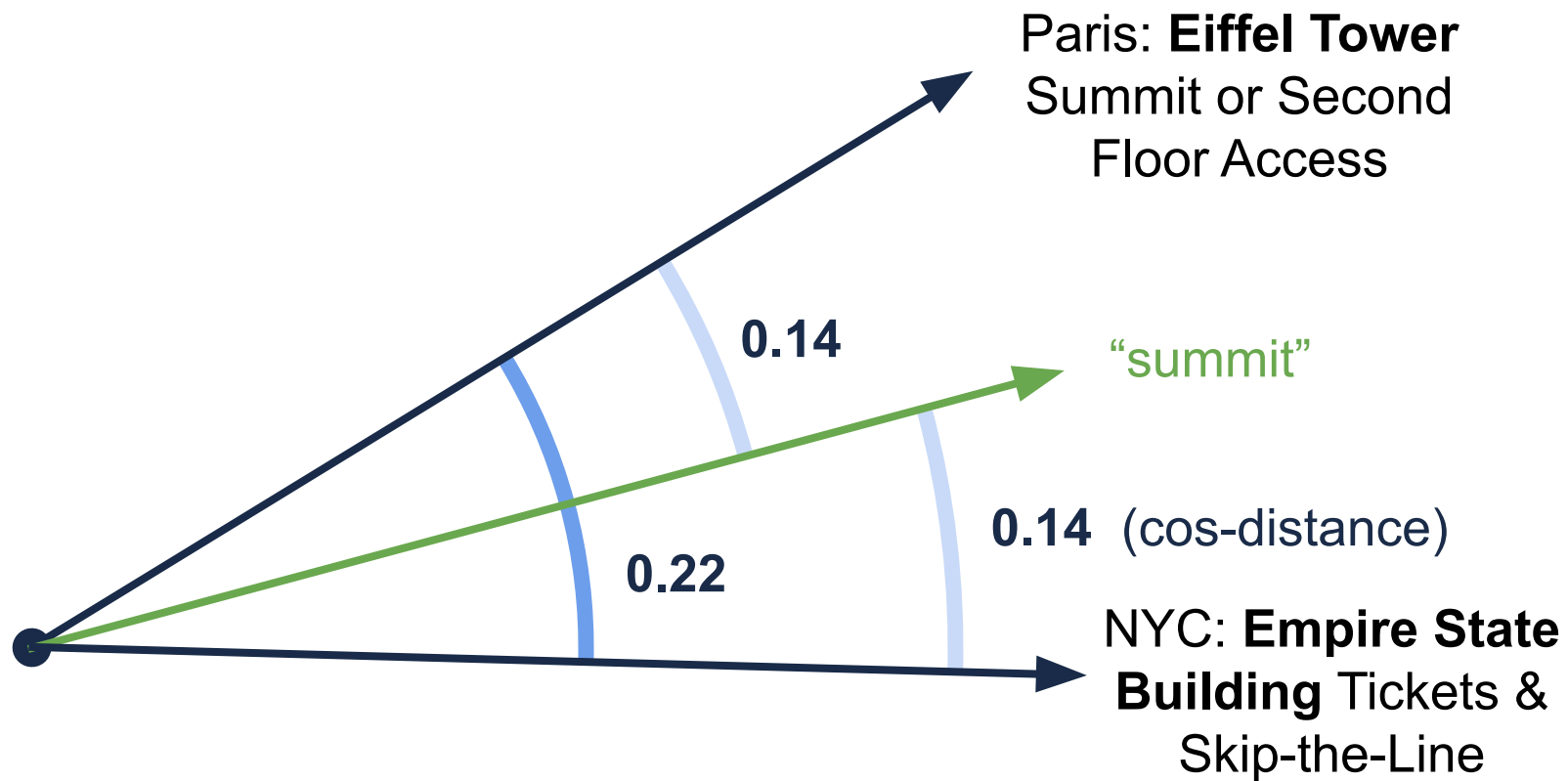




# Semantic Vectors at GetYourGuide: Example

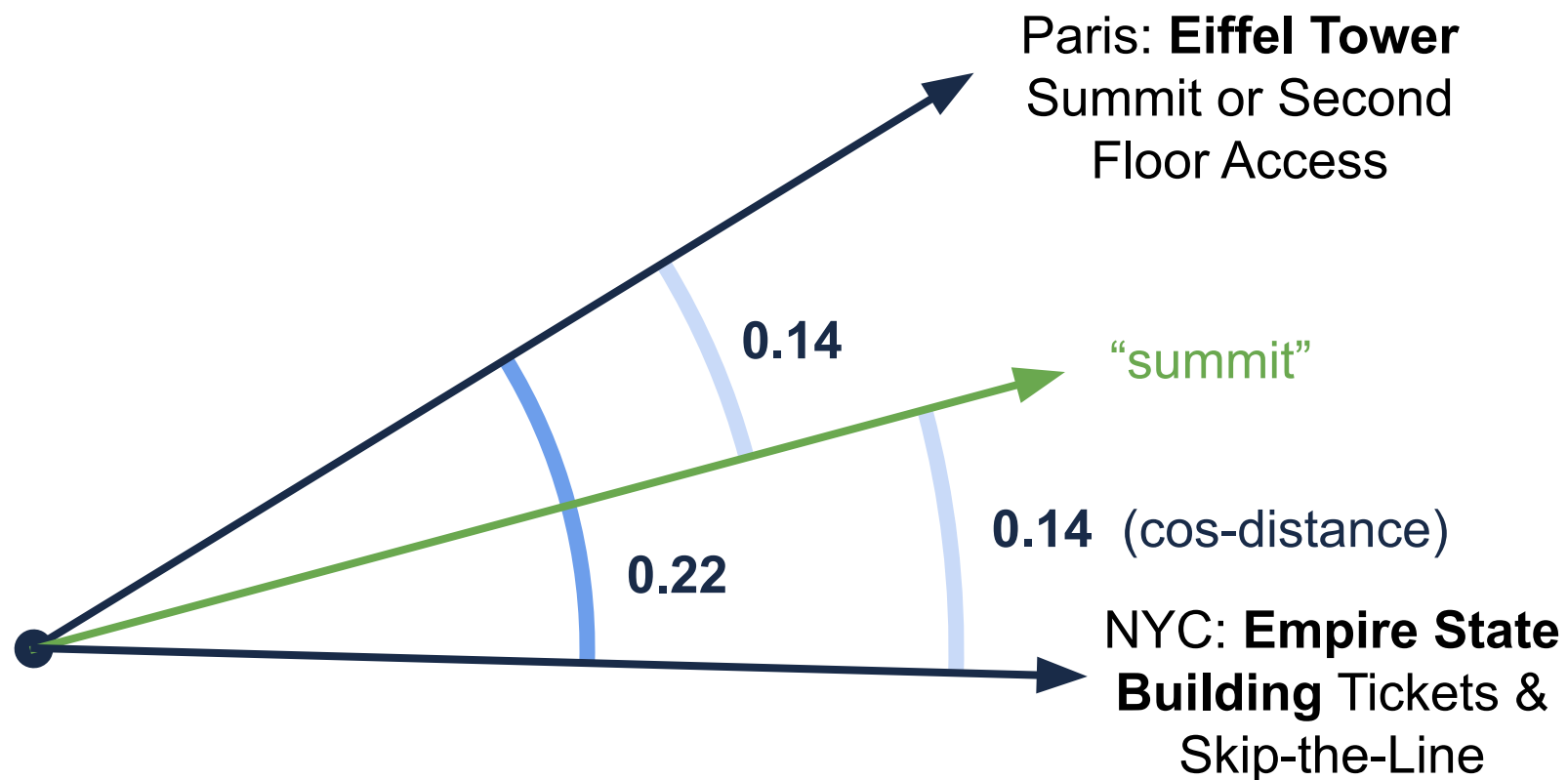


# Semantic Vectors at GetYourGuide: Example



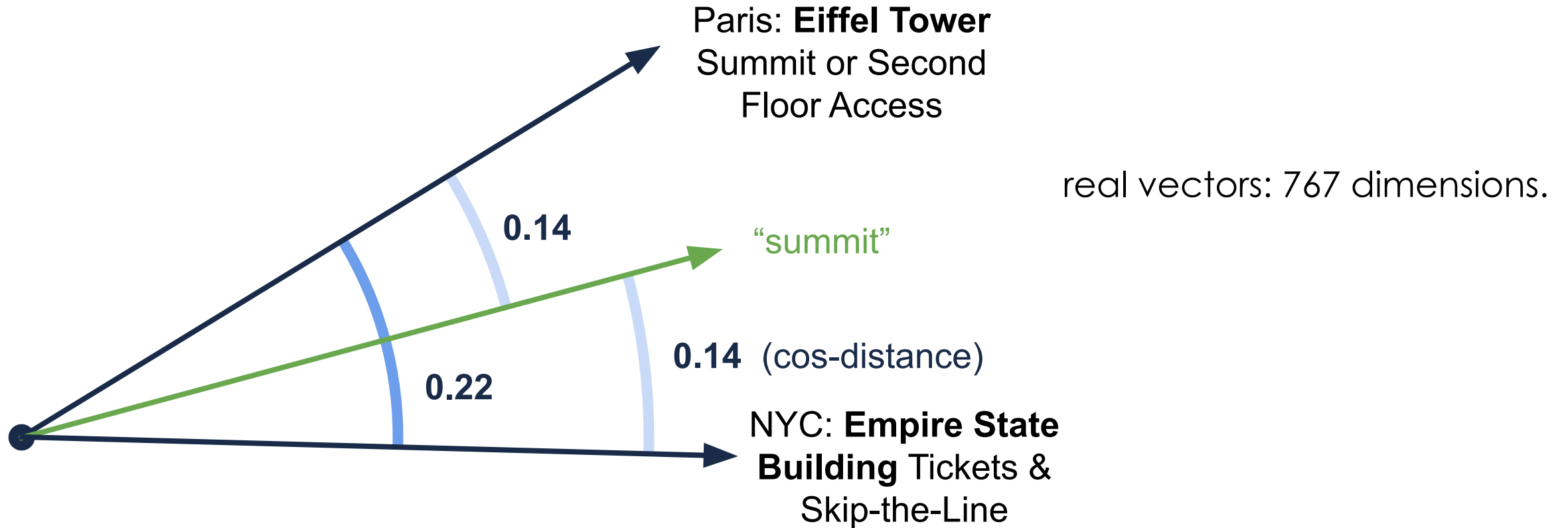


## Semantic Vectors at GetYourGuide: Example



**How far is the Empire State Building from the Eiffel Tower? 0.22**  
... and both are 0.14 from the keyword "summit".

# Semantic Vectors at GetYourGuide: Example



**How far is the Empire State Building from the Eiffel Tower? 0.22**  
... and both are 0.14 from the keyword "summit".





# Learnings over Time



GET YOUR GUIDE

# ≤ 2018: LDA for Similar Recommender

- *application*: recommendations of similar activities
- *vectors for*: activities
- *model*: [LDA: Latent Dirichlet allocation](#) (2003)
  - statistical model on word frequencies in documents
  - automatically discovers K topics (K: input value)
  - output: document → K-dimensional vector of probabilities
  - ⇒ coordinates have clear meaning
- **learning: start simple**
  - or: even hand-designed vectors!

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple



# 2018: Word2Vec for Similar Recommender

- *application*: recommendations of similar activities
- *vectors for*: activities
- *model*: [Word2Vec](#) (2013)
  - 2-layer neural network trained on documents
  - output: word  $\rightarrow$  vector
- **learning**: ~~text~~  $\rightarrow$  **user journey** <sup>1</sup>
  - word  $\rightarrow$  page of web site
  - document  $\rightarrow$  observed user journey (sequence of pages)
  - $\Rightarrow$  vectors for pages
  - $\Rightarrow$  vector for activities

<sup>1</sup> idea presented [by AirBnB at KDD 2018](#)

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey

# 2021: Word2Vec for Ranking on Landing Pages

- *application*: activity-based landing page
  - users arrive with interest in specific activity  $s$
  - rank rest of activities  $a_i$  on landing page by:  
$$(1 - w)\text{ranking\_score}(a_i) + w \text{vector\_sim}(a_i, s)$$
- *vectors for*: activities
- *model*: [Word2Vec](#) (2013) (re-used!)
- **learnings**:
  - **re-use vectors**
  - **do regular, cross-team hackathons**



# 2021: Word2Vec for Ranking on Landing Pages

- *application*: activity-based landing page
  - users arrive with interest in specific activity  $s$
  - rank rest of activities  $a_i$  on landing page by:  
$$(1 - w)\text{ranking\_score}(a_i) + w \text{ vector\_sim}(a_i, s)$$
- *vectors for*: activities
- *model*: [Word2Vec](#) (2013) (re-used!)
- **learnings**:
  - **re-use vectors**
  - **do regular, cross-team hackathons**

with:



**Viktoriia Kucherenko**  
Software Engineer



**Ashraf Aaref**  
Software Engineer

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors, hackathons

# 2022: Word2Vec for Ad Suggestion Diversity

- *application*: text suggestion diversity for ads
  - we provide headline and description suggestions
  - external algorithm tries combinations
  - we:
    1. removed too similar candidates
    2. ranked suggestions iteratively, pick  $s$  maximizing [MMR](#):
$$(1-w)\text{relevance}(s) + w \min_{a \in \text{ranked\_above}} \text{dist}(s, a)$$
- *vectors for*: headlines and descriptions
- *model*: [Word2Vec](#) (2013) - this time applied to text!
- **learnings: look at (bad performing) examples!**
  - revealed undesired dist values  $\Rightarrow$  removal of stop words



# 2022: Word2Vec for Ad Suggestion Diversity



**Marina Zemskova**  
Data Scientist

- *application*: text suggestion diversity for ads
  - we provide headline and description suggestions
  - external algorithm tries combinations
  - we:
    1. removed too similar candidates
    2. ranked suggestions iteratively, pick  $s$  maximizing MMR:
$$(1-w)\text{relevance}(s) + w \min_{a \in \text{ranked\_above}} \text{dist}(s, a)$$
- *vectors for*: headlines and descriptions
- *model*: Word2Vec (2013) - this time applied to text!
- **learnings: look at (bad performing) examples!**
  - revealed undesired dist values  $\Rightarrow$  removal of stop words

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors
2022	Word2Vec	Text Suggestion Diversity for Ads	Clicks: +4%	look at (bad performing) examples

# 2022: DistilBERT for Cross-Sell Recommender

- *application*: cross-sell recommendations
  - after booking activity  $b$  you might also like  $a$
  - LightGBM model
  - use  $b$ 's and  $a$ 's vectors as input
- *vectors for*: activities
- *model*: [DistilBERT](#) (2019)
- **learnings**:
  - **fine-tune LLMs** on own data
  - **re-use models**
    - here: from internal keyword predictions
  - **decrease model complexity**
    - here: light DistilBERT model + PCA to 3 dimensions



# 2022: DistilBERT for Cross-Sell Recommender

- *application*: cross-sell recommendations
  - after booking activity  $b$  you might also like  $a$
  - LightGBM model
  - use  $b$ 's and  $a$ 's vectors as input
- *vectors for*: activities
- *model*: [DistilBERT](#) (2019)
- **learnings**:
  - **fine-tune LLMs** on own data
  - **re-use models**
    - here: from internal keyword predictions
  - **decrease model complexity**
    - here: light DistilBERT model + PCA to 3 dimensions



**Stephane Leonard**  
Data Scientist

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors
2022	Word2Vec	Text Suggestion Diversity for Ads	Clicks: +4%	look at (bad performing) examples
2022	DistilBERT	Cross-Sell Recom.	CR: +6% <sup>1</sup>	fine-tune LLM, re-use models, decrease model complexity

<sup>1</sup> tested together with other changes

# 2022: TourBERT for External Search Queries

- *application*: ranking for external search queries  
(on landing pages based on Google ad keyword)  
e.g. “sunset cruise Airlie Beach”



sunset cruise airlie beach

- Images
- Maps
- Shopping
- News
- Videos
- Books
- Flights
- Finance

About 1,100,000 results (0.48 seconds)

Results for **Airlie Beach QLD, Australia** · Choose area

**Sponsored**

 GetYourGuide  
<https://www.getyourguide.com> > airliebeach > boat

**Sunset Cruise Airlie Beach - Book Top Boat & River Tour**

See your destination from the water. Find top-rated boat tours and **crui**ses. Discover and book over 40,000 of the best things to do around the world on GetYourGuide. Free Cancellation. Best value guarantee. Multiple languages. Flexible booking. Expert guides.



Airlie Beach

# Cruises & boat tours

OLD:



DAY TRIP

### Airlie Beach: Whitsundays and Whitehaven Half-Day Cruise

🕒 Duration: 6 hours

A white catamaran sailing on blue water with a green forested coastline in the background. A green globe icon and a heart icon are overlaid on the top left and top right of the image respectively.

**WATER ACTIVITY**

From Airlie: Whitsundays Full-Day Catamaran Sailing Trip

🕒 Duration: 10 hours

★★★★★ 4.7 (374 reviews)

**From € 140.20** per person

An aerial view of a yellow glass-bottom boat on a shallow reef flat. A green globe icon and a heart icon are overlaid on the top left and top right of the image respectively.

**WATER ACTIVITY**

Airlie Beach: Glass-Bottom Boat Tour

🕒 Duration: 1 hour

★★★★★ 4.4 (52 reviews)

**From € 52.99** per person

A red and white sailboat on blue water. A green globe icon and a heart icon are overlaid on the top left and top right of the image respectively.

**WATER ACTIVITY**

Tongarra: All-Inclusive Day Sail

🕒 Duration: 8 hours

★★★★★ 5.0 (13 reviews)

**From € 99.95** per person

A sailboat on blue water with a green forested coastline in the background. A green globe icon and a heart icon are overlaid on the top left and top right of the image respectively.

**MULTI-DAY TRIP**

From Airlie Beach: 2-Day Whitsundays Sailing & Camping Trip

🕒 Duration: 2 days

★★★★★ 5.0 (1 Review)

**From € 214.66** per person

GET YOUR GUIDE

Airlie Beach

# Cruises & boat tours

**NEW:**



WATER ACTIVITY

## Airlie Beach: 2-Hour Sunset Cruise with Sparkling Wine

🕒 Duration: 2 hours



WATER ACTIVITY

## Airlie Beach: Sunset Sail to Whitsundays

🕒 Duration: 2 hours

★★★★★ 4.7 (74 reviews)  
**From € 52.99** per person



WATER ACTIVITY

## Airlie: Sunset Sail with Aperol Spritz and Antipasto

🕒 Duration: 2 hours

★★★★★ 5.0 (5 reviews)  
**From € 66.41** per person



DAY TRIP

## Airlie Beach: Whitsundays and Whitehaven Half-Day Cruise

🕒 Duration: 6 hours

★★★★★ 4.4 (348 reviews)  
**From € 80.50** per person



WATER ACTIVITY

## Airlie Beach: Full-Day Whitehaven Beach & Islands Boat Tour

🚗 Pickup available

🕒 Duration: 1 day

★★★★★ 5.0 (4 reviews)  
**From € 117.39** per person

**GET YOUR GUIDE**

# 2022: TourBERT for External Search Queries

- *application*: ranking for external search queries  
(on landing pages based on Google ad keyword)  
e.g. “sunset cruise Airlie Beach”

# 2022: TourBERT for External Search Queries

- *application*: ranking for external search queries (on landing pages based on Google ad keyword)
  - rank activities  $a_i$  for query (keyword)  $q$  by:  
$$(1 - w)\text{ranking\_score}(a_i) + w \text{vector\_sim}(a_i, q)$$
- *vectors for*: activities, queries (ad keywords)
- *model*: [TourBERT](#) (2022) ( $\rightarrow$  vectors from title)
- **learnings**:
  - **vector arithmetics are useful**: here:
    - subtract location part (avg. vector of all activities from location) and ...
    - simple query encoding: avg. vector of best-performing activities (like: [Mandal, MICES 2021](#))
  - **re-assess solutions**



# 2022: TourBERT for External Search Queries

with:



**Ryan Sequeira**  
Data Scientist



**Hsin-Ting Hsien**  
MLOps Engineer

- *application*: ranking for external search queries (on landing pages based on Google ad keyword)
  - rank activities  $a_i$  for query (keyword)  $q$  by:  
$$(1 - w)\text{ranking\_score}(a_i) + w \text{vector\_sim}(a_i, q)$$
- *vectors for*: activities, queries (ad keywords)
- *model*: [TourBERT](#) (2022) ( $\rightarrow$  vectors from title)
- **learnings**:
  - **vector arithmetics are useful**: here:
    - subtract location part (avg. vector of all activities from location) and ...
    - simple query encoding: avg. vector of best-performing activities (like: [Mandal, MICES 2021](#))
  - **re-assess solutions**

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors
2022	Word2Vec	Text Suggestion Diversity for Ads	Clicks: +4%	look at (bad performing) examples
2022	DistilBERT	Cross-Sell Recom.	CR: +6% <sup>1</sup>	fine-tune LLM, re-use models, decrease model complexity
2022	TourBERT	Ranking for External Search Queries	CR: +2% <sup>2</sup> (+2% <sup>3</sup> )	vector arithmetic, simple query encoding, re-assess solutions

<sup>1</sup> tested together with other changes

<sup>2</sup> only on specific keywords

<sup>3</sup> by removing diversity

# 2022: Single Feature for Ranking Diversity

- *application*: increase diversity of ranked activities
  - rank activities  $a_i$  by:  
$$(1-w)\text{ranking\_score}(a_i) + w \text{ avg}_{a \in \text{ranked\_above}} \text{dist}(a_i, a)$$
- *vectors for*: activities
- *model*: ~~FourBERT~~ (2022) → 1 feature
- **learnings**:
  - **start simple** (learned again 😊)
    - ~~vector distance~~ → share of newly covered categories by new activity

# 2022: Single Feature for Ranking Diversity

- *application*: increase diversity of ranked activities
  - rank activities  $a_i$  by:  
 $(1-w)\text{ranking\_score}(a_i) + w \text{ avg}_{a \in \text{ranked\_above}} \text{dist}(a_i, a)$
- *vectors for*: activities
- *model*: ~~FourBERT~~ (2022) → 1 feature
- **learnings**:
  - **start simple** (learned again 😏)
    - ~~vector distance~~ → share of newly covered categories by new activity



**Galina Baimeeva**  
Data Scientist



**Maximilian Jenders**  
Data Scientist



# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors
2022	Word2Vec	Text Suggestion Diversity for Ads	Clicks: +4%	look at (bad performing) examples
2022	DistilBERT	Cross-Sell Recom.	CR: +6% <sup>1</sup>	fine-tune LLM, re-use models, decrease model complexity
2022	TourBERT	Ranking for External Search Queries	CR: +2% <sup>2</sup> (+2% <sup>3</sup> )	vector arithmetic, simple query encoding, re-assess solutions
2022	<del>TourBERT</del> → 1 Value	Diversity in Activity Ranking	CR: +1% <sup>4</sup>	start simple (learned again 😊)

<sup>1</sup> tested together with other changes    <sup>2</sup> only on specific keywords    <sup>3</sup> by removing diversity    <sup>4</sup> ongoing experiment, not (yet?) significant

# Learnings over Time

CR = conversion rate

Started	Model	Usage	Uplift?	Learnings
≤ 2018	LDA	Similar Activity Recom.	CR: +	start simple
2018	Word2Vec	Similar Activity Recom.	CR: +	<del>text</del> → user journey
2021	Word2Vec	Activity-Based Landing Pages	CR: +4%	re-use vectors
2022	Word2Vec	Text Suggestion Diversity for Ads	Clicks: +4%	look at (bad performing) examples
2022	DistilBERT	Cross-Sell Recom.	CR: +6% <sup>1</sup>	fine-tune LLM, re-use models, decrease model complexity
2022	TourBERT	Ranking for External Search Queries	CR: +2% <sup>2</sup> (+2% <sup>3</sup> )	vector arithmetic, simple query encoding, re-assess solutions
2022	<del>TourBERT</del> → 1 Value	Diversity in Activity Ranking	CR: +1% <sup>4</sup>	start simple (learned again 😊)

<sup>1</sup> tested together with other changes    <sup>2</sup> only on specific keywords    <sup>3</sup> by removing diversity    <sup>4</sup> ongoing experiment, not (yet?) significant



# Top Learnings



GET YOUR GUIDE

## Top-6 Learnings

1. start simple / decrease model complexity
2. look at (bad performing) examples
3. re-use models and vectors
4. use vector arithmetics
5. re-assess solutions
6. do regular, cross-team hackathons





# Outlook

GET YOUR GUIDE

# Outlook: Semantic Vectors at GetYourGuide

- other use cases:
  - internal search
  - matching categories ↔ activities
  - chatbot attempts,
  - ...
- using new / better models:
  - dense encoders for query and document (Siamese → two-tower)
  - newest generation of LLMs

# Outlook: Semantic Vectors at GetYourGuide

**We are hiring**

[getyourguide.careers](https://getyourguide.com/careers)

- other use cases:
  - internal search
  - matching categories ↔ activities
  - chatbot attempts,
  - ...
- using new / better models:
  - dense encoders for query and document (Siamese → two-tower)
  - newest generation of LLMs



**We are hiring**

[getyourguide.com/careers](https://getyourguide.com/careers)

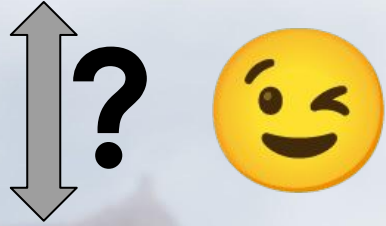
**Thank you!**

**GET YOUR GUIDE**

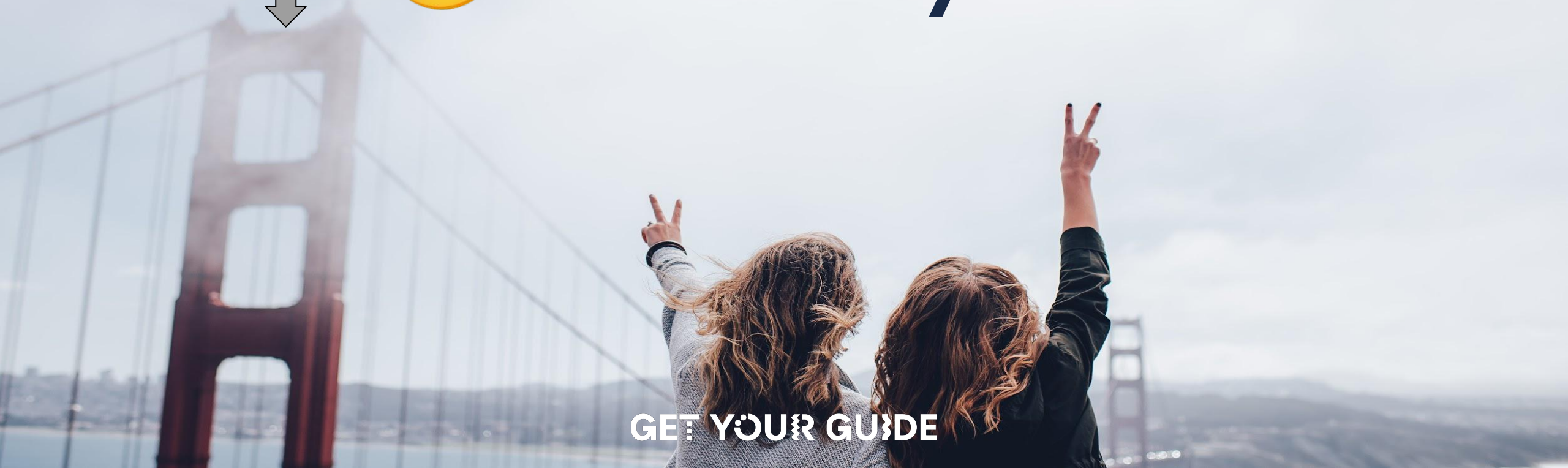




**We are hiring**  
[getyourguide.careers](https://getyourguide.com/careers)



**Thank you!**



**GET YOUR GUIDE**