The background of the slide features a dark navy blue field with a large, dynamic burst of red particles on the left side. These particles, which vary in size and opacity, appear to be exploding or dispersing from the left edge towards the center, creating a sense of movement and intensity. The overall effect is reminiscent of a microscopic view of a virus or a high-speed camera shot of a powder explosion.

Disease Spread Predictions

By: Billy Adams

Problem Statement

- Highlight the lack of a quality global disease surveillance system
- Motivation for improved and better-defined anti-poaching control measures

Business Impact

- Better preparation for healthcare systems
- Better traceability of diseases
- Better allocation of resources
- Precision for more funding

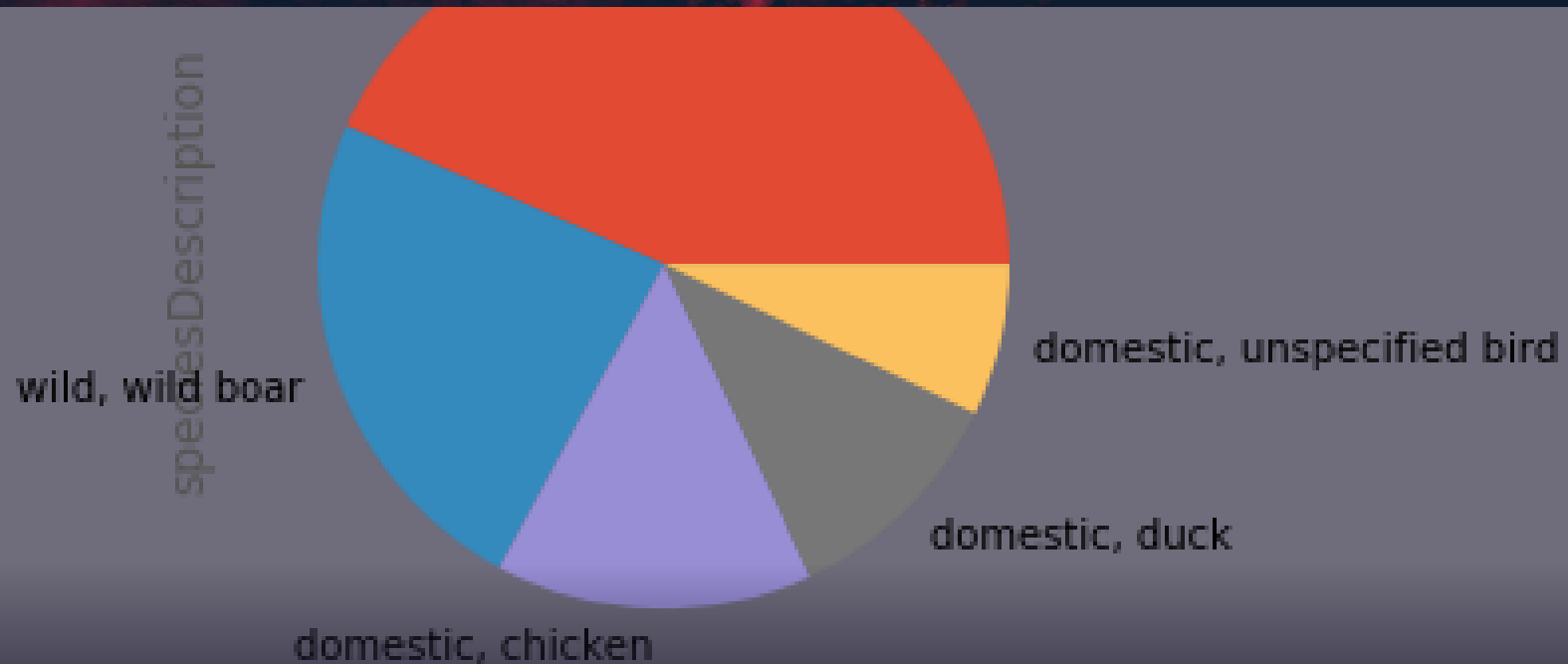
Methodology

- EDA of EMPRES Global Animal Disease Surveillance dataset
- Feature engineering
- SARIMAX
- LSTM
- EDA of CITES dataset
- SIR Model

Methodology (cont.)

EDA of EMPRES Global
Animal Disease
Surveillance dataset

What species is most likely to
cause an outbreak?

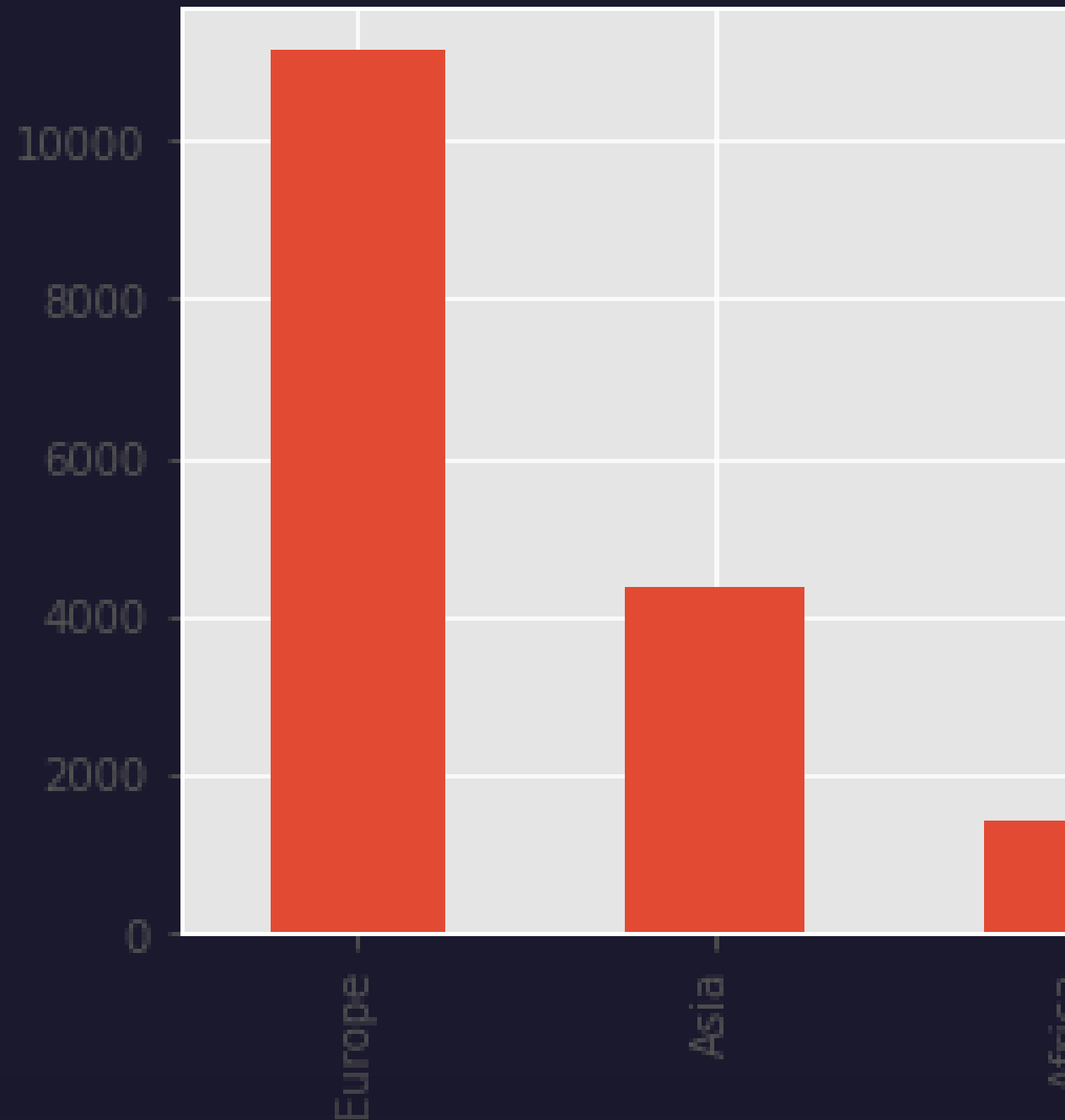


Methodology (cont.)

EDA of EMPRES

Global Animal Disease
Surveillance dataset

What region/country is most at
risk according to this data?



Methodology (cont.)

Feature engineering

Does the difference between the reported date and the reported date have any relationship to total deaths and can the sum at risk be used to predict cases?

Days_since_obs	porportion_of_cases_to_at_risk_pop	sumCases	sumDeaths	
Days_since_obs	1.000000	-0.272359	0.011435	0.020488
porportion_of_cases_to_at_risk_pop	-0.272359	1.000000	0.092315	0.148940
sumCases	0.011435	0.092315	1.000000	0.997713
sumDeaths	0.020488	0.148940	0.997713	1.000000

Methodology (cont.)

SARIMAX

MAE: 697.401

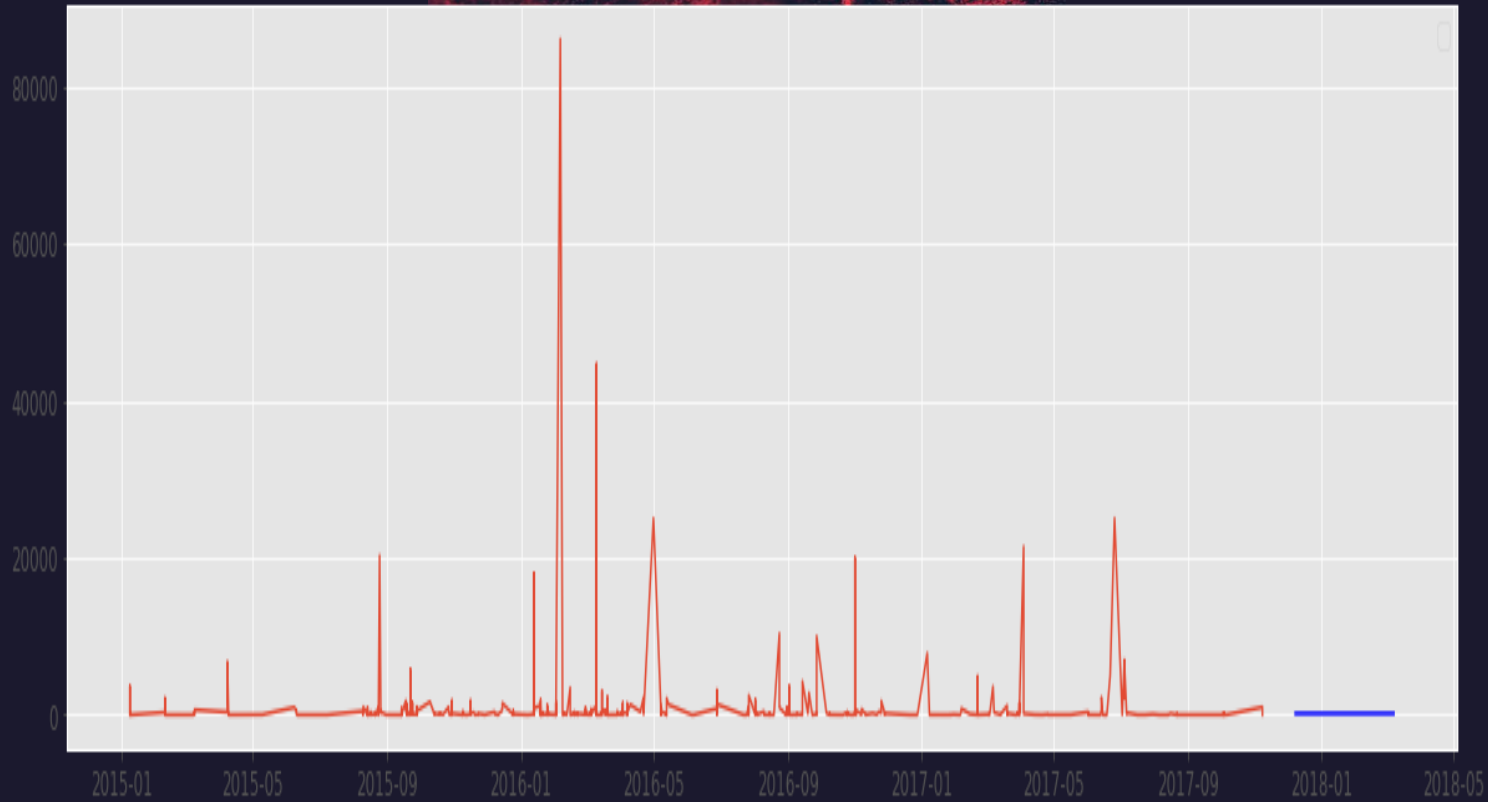
Test set statistics

sumCases

count	12.000000
mean	121.000000
std	275.234115
min	2.000000
25%	8.250000
50%	31.500000
75%	53.000000
max	975.000000

Predictions statistics

```
count
12.000000
mean
749.601185
std
388.937636
min
337.154135
25%
473.952395
50%
558.427543
75%
1004.27350
5 max
1506.45808
1
```

Methodology (cont.) LSTM

Methodology (cont.)

EDA on cites dataset



Comparisons of wildlife exports and disease outbreaks

Top 5 countries of origin for trade species

ID -Indonesia

XX - Unknown

US - United States of America

VN - Vietnam

ZW - Zimbabwe

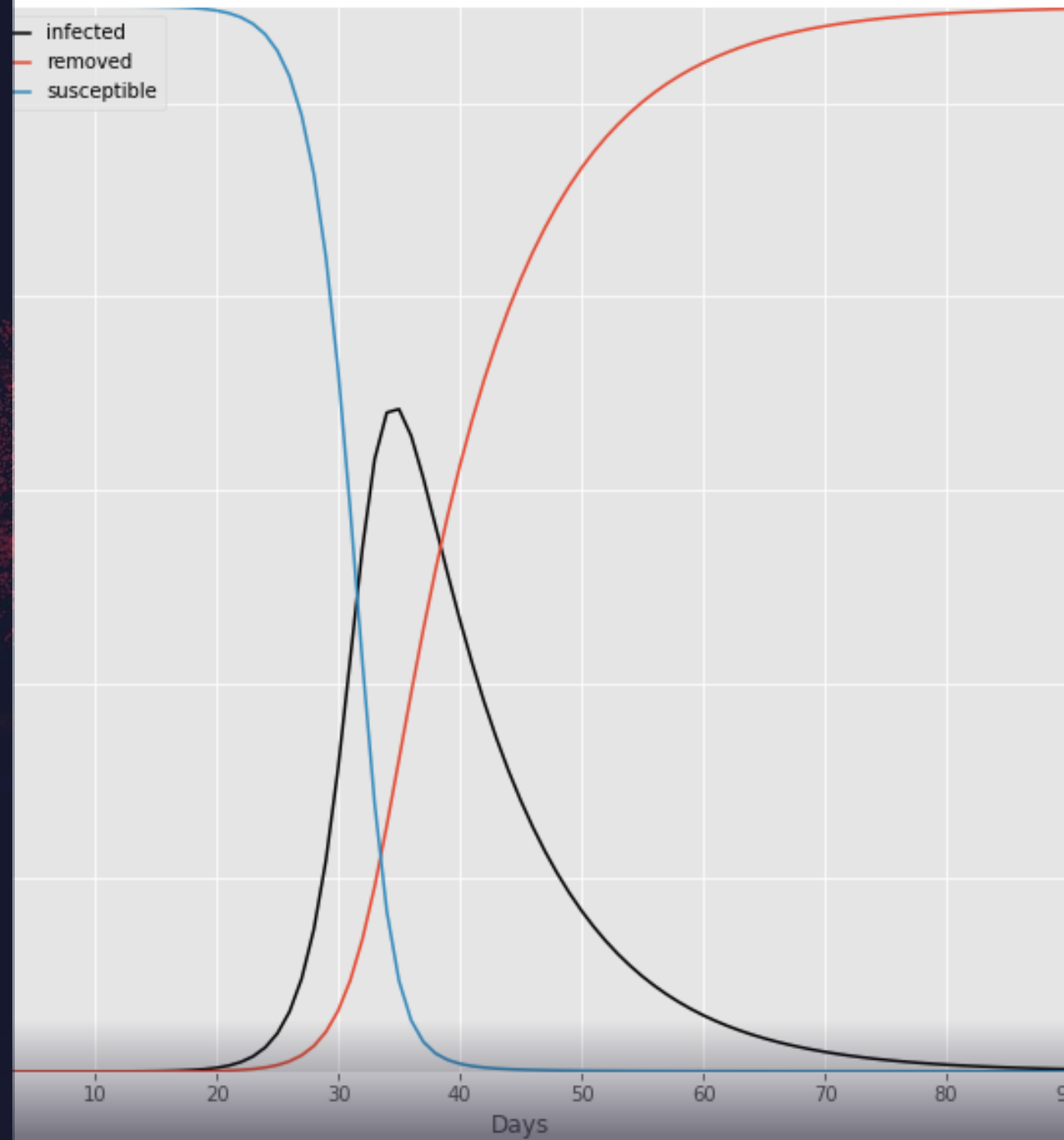


Exotic species of interest

- Elephantidae spp. - Asian elephant
- Loxodonta africana - African bush elephant
- Macaca fascicularis - Long-tailed macaque
- Varanus salvator - Asian water monitor
- Python bivittatus - Burmese python
- Crocodylus porosus - Saltwater crocodile



Methodology (cont.) SIR Model

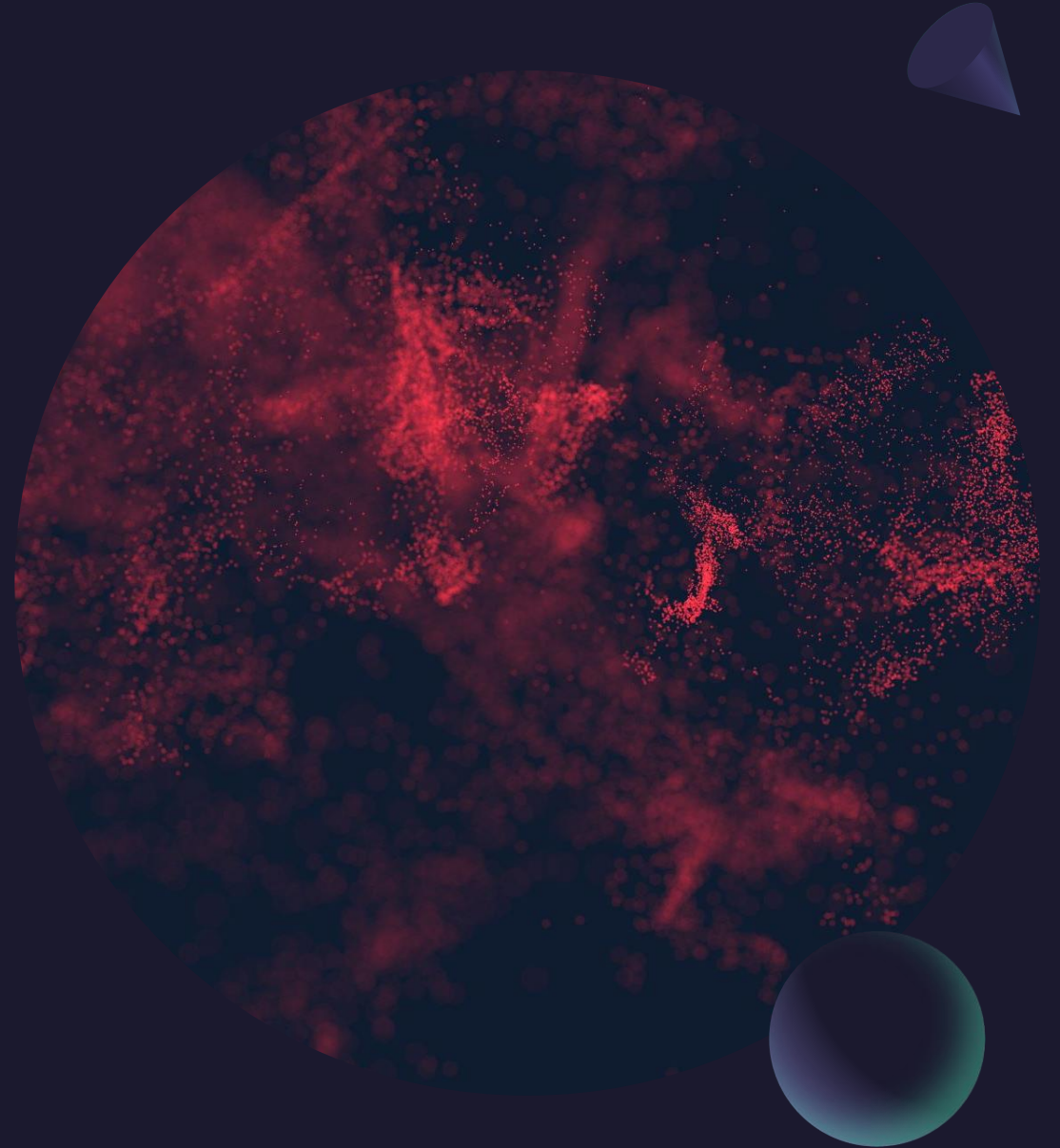


Findings

Many factors influence disease spread and outbreaks

Qualified and trained staff are needed to sustain a global disease surveillance system

Better anti-poaching measures can help mitigate the emergence of the diseases



Future work

Simulation of disease spread with agents such as weather, population, control measures, etc.

Image classification of animals with the disease and without

Gather data on the healthcare systems to see where more of our attention should be



THANK YOU

The background features a dark blue gradient with dynamic, abstract patterns of red and orange particles. These particles are concentrated in the lower right and center, forming swirling, cloud-like shapes that suggest movement and energy. The overall effect is a modern, artistic backdrop for the text.