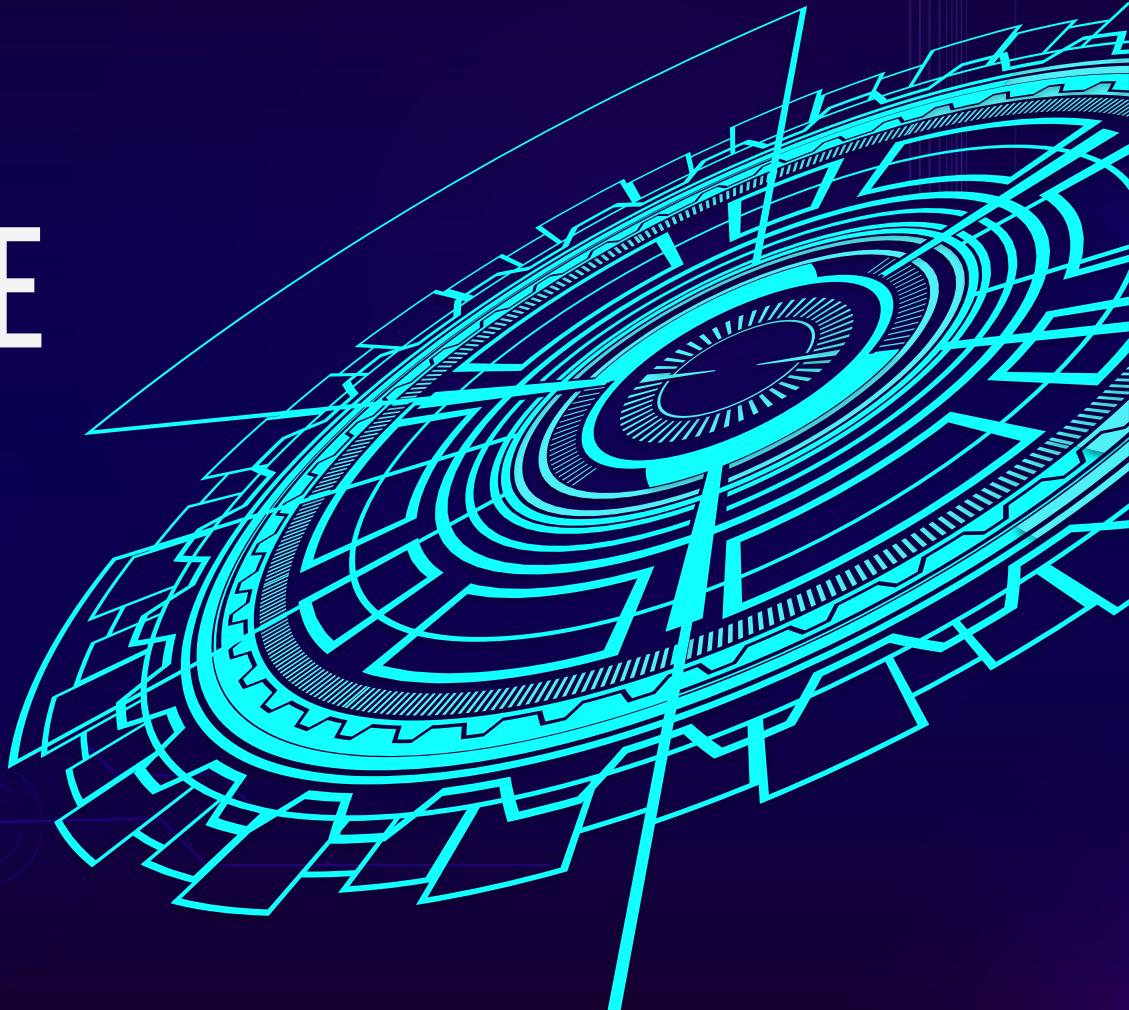




FORMULA E STRATEGY



OUR TEAM

MARYAM
SOOMRO



MUNISH
KHURANA



ISIOMA OCHIA



ZHUXUAN
(SHAWN) XU

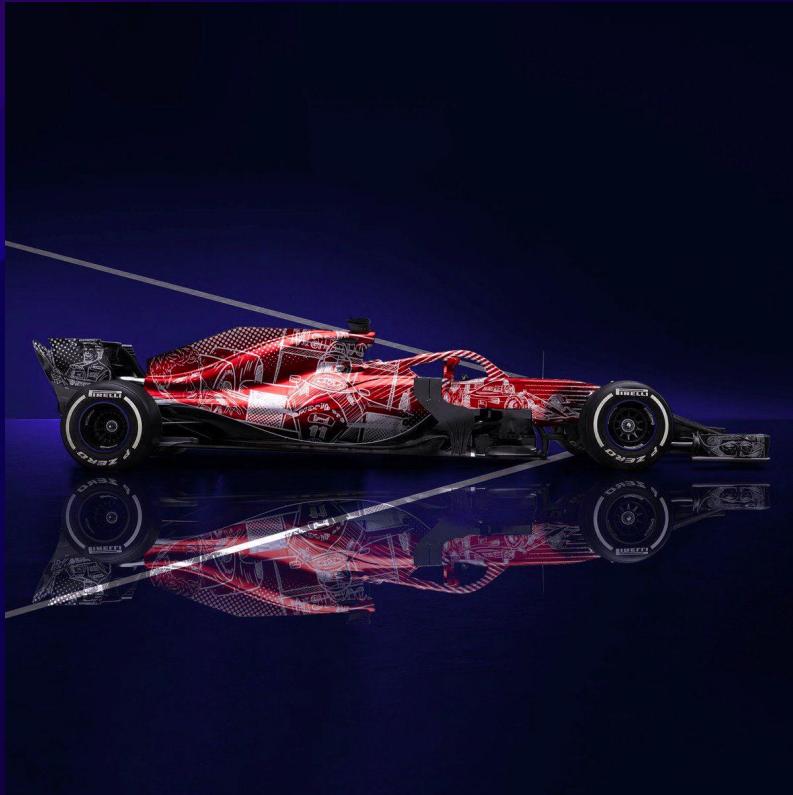


ELLEN ZHANG



JIANGKUN
XIONG





AGENDA

01

Client Background

02

Problem Overview

03

Models and Analysis

04

Recommendations

A Formula 1 race car is shown from a front-three-quarter perspective, driving on a track at night. The car's headlights illuminate the road ahead, creating bright streaks of light against the dark background. The track surface shows signs of wear and tire marks. The overall atmosphere is dynamic and focused.

01 Case Background



“Formula E is a single-seater motorsport championship that uses only **electric cars**. The series is promoted and owned by Formula E Holdings”



45min



1 lap



12 teams



1. Practice

- Capture data for the race
- No points awarded

2. Qualifying & Superpole

- start position
- fastest driver

3. Race

- 25pt 1st place
- 18pt 2nd place
- Fastest Lap



Formula E Racing



Disrupting
Motor
Racing



Fully Electric



Sustainable
Environmentally
Friendly



Background | Problem | Models and Analysis | Recommendation

02 Problems





Challenge Facing



Weather



Skill

With similar electric battery - powered cars,

how can a team realize advantage in a race?

Background | Problem | Models and Analysis | Recommendation



THE ISSUE

How to use collected data to predict the number of laps remained and compete strategically in the final race?

03 Models & Analysis





MODELING PROCESS



DATA PREPARATION

Merge data and
clean files

MODEL SELECTION

Random Forest
Regression

TRAINING & VALIDATION

Fit data to model

ANALYSIS & RESULTS

Accuracy & implication





MODELING PROCESS



DATA PREPARATION

Merge data and clean files

MODEL SELECTION

Random Forest
Regression

TRAINING & VALIDATION

Fit data to model



ANALYSIS & RESULTS

Accuracy & implication





Data



Weather Data

wind speed
air temperature
wind speed
wind direction



Race Data

lap time
lap number
KPH
lap improvements



Cleaned Dataset

bridge time data from
race data and weather
data



Data



Weather Data

wind speed
air temperature
wind speed
wind direction

Race Data

lap time
lap number
KPH
lap improvements

Cleaned Dataset

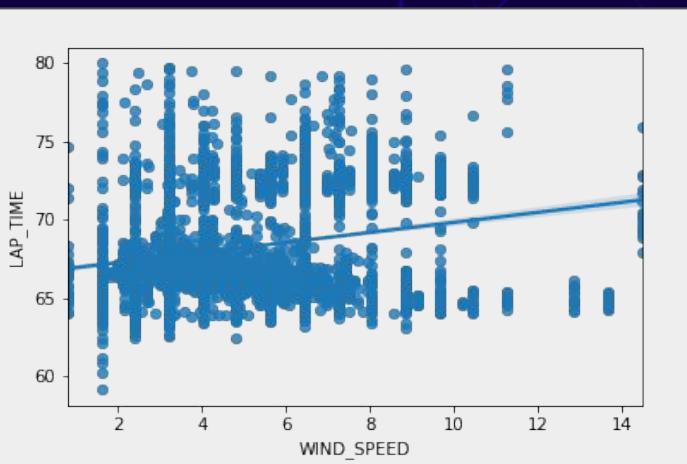
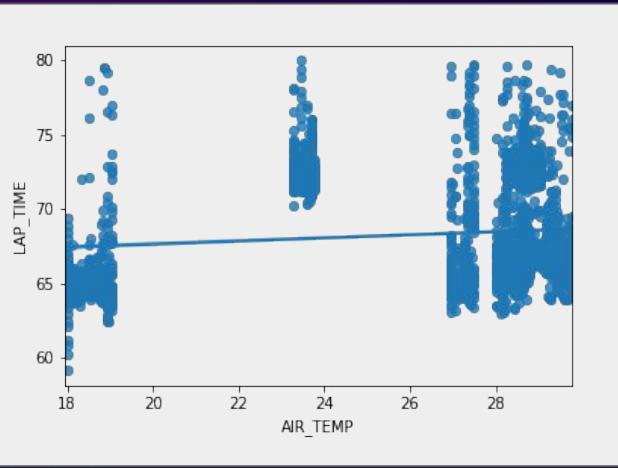
bridge time data from
race data and weather
data



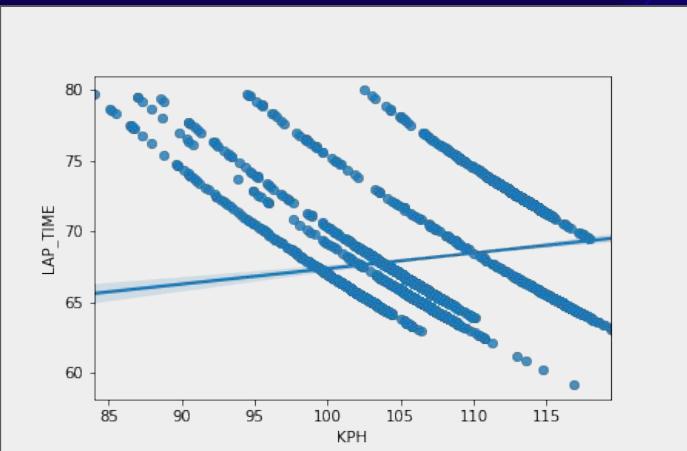
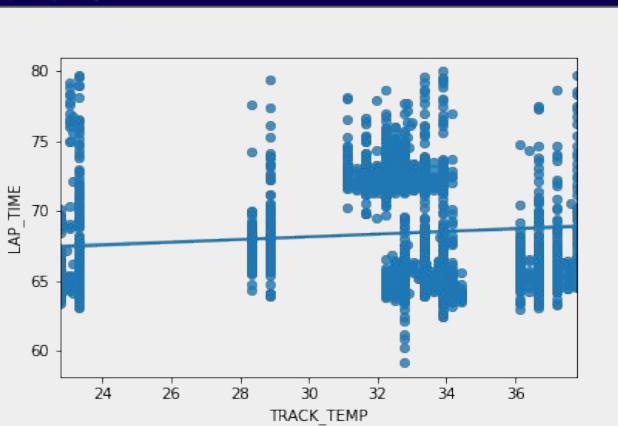
	LAP_TIME	TRACK_TEMP	AIR_TEMP	WIND_SPEED	WIND_DIRECTION			
2	68.321	39.44445	20.3611	8.04672	311.5			
5	69.630	38.88890	20.3889	14.48410	277.0			
8	NUMBER	LAP_NUMBER	LAP_TIME	LAP_IMPROVEMENT	KPH	ELAPSED	HOUR	DRIVER_NAME
11		2	3	68.321	4	120.0	6:25.812	2020-11-10 10:00:05.812 Sam BIRD
14	GROUP	TEAM	MANUFACTURER	AIR_TEMP	TRACK_TEMP	WIND_SPEED	WIND_DIRECTION	
27	1	DS Virgin Racing	VIRGIN	20.3611	39.44445	8.04672	311.5	
28	1	Andretti Formula E	ANDRETTI	20.3889	38.88890	14.48410	277.0	
47	1	Andretti Formula E	ANDRETTI	20.3611	39.44445	8.04672	311.5	
66	1	Panasonic Jaguar Racing	JAGUAR	20.3889	38.88890	14.48410	277.0	
	1	ABT Schaeffler Audi Sport	ABT	20.3889	38.88890	14.48410	277.0	



Background | Problem | **Models and Analysis** | Recommendation



Visualizations





MODELING PROCESS





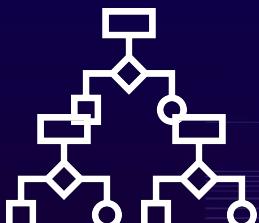
RANDOM FOREST

- Equally leverages all predictive features
- Avoids overfitting
- Computationally lenient



NEURAL NETWORK

- Complex modelling
- Computationally expensive
- Results are difficult to interpret





MODELING PROCESS





RANDOM FOREST



```
pred_res = []

for i in range(500):
    totaltime = 0
    numlaps = 0
    i = 0

    while totaltime <= 45:
        #Input ranges are derived from minimum and maximum values from each
        #category in the dataset
        input = {'AIR_TEMP' : [(random.randrange(17, 29))],
                 'TRACK_TEMP': [(random.randrange(22,37))],
                 'WIND_SPEED': [(random.randrange(0,14))],
                 'ELAPSED': [(random.randrange(59,3459))],
                 'KPH': [(random.randrange(84,119))]}
        print(input)
        input = pd.DataFrame(data=input)
        y_pred = (regr.predict(input))/60
        totaltime += y_pred
        numlaps += i
        i += 1
    pred_res.append(i)
```



MODELING PROCESS





NUMBER OF LAPS SIMULATION DISTRIBUTION

Trials: 500

Minimum Predicted Laps: 39.00

Maximum Predicted Laps: 41.00

Average Predicated Laps: 39.86

St. Dev of Predicated Laps:
0.59

$R^2 \sim 0.97$

MSE ~0.0004

TRAINING REPORTS

04 Recommendations





RANDOM FOREST REGRESSION



Refined data to increase
the value of unstructured
data for prediction

DATA VALUE



MODEL FLEXIBILITY

Improve models accuracy
based on different race and
external factors



REAL TIME SIMULATION



```
pred_res = []

for i in range(500):
    totaltime = 0
    numlaps = 0
    i = 0

    while totaltime <= 45:
        #Input ranges are derived from minimum and maximum values from each
        #category in the dataset
        input = {'AIR_TEMP' : [(random.randrange(17, 29))],
                 'TRACK_TEMP': [(random.randrange(22,37))],
                 'WIND_SPEED': [(random.randrange(0,14))],
                 'ELAPSED': [(random.randrange(59,79))],
                 'KPH': [(random.randrange(84,119))]}
        print(input)
        input = pd.DataFrame(data=input)
        y_pred = (regr.predict(input))/60
        totaltime += y_pred
        numlaps += i
        i += 1
    pred_res.append(i)
```



SWOT ANALYSIS

STRENGTHS

Rich data volume
Real-time simulation

THREATS

High competition in the
data analytics landscape

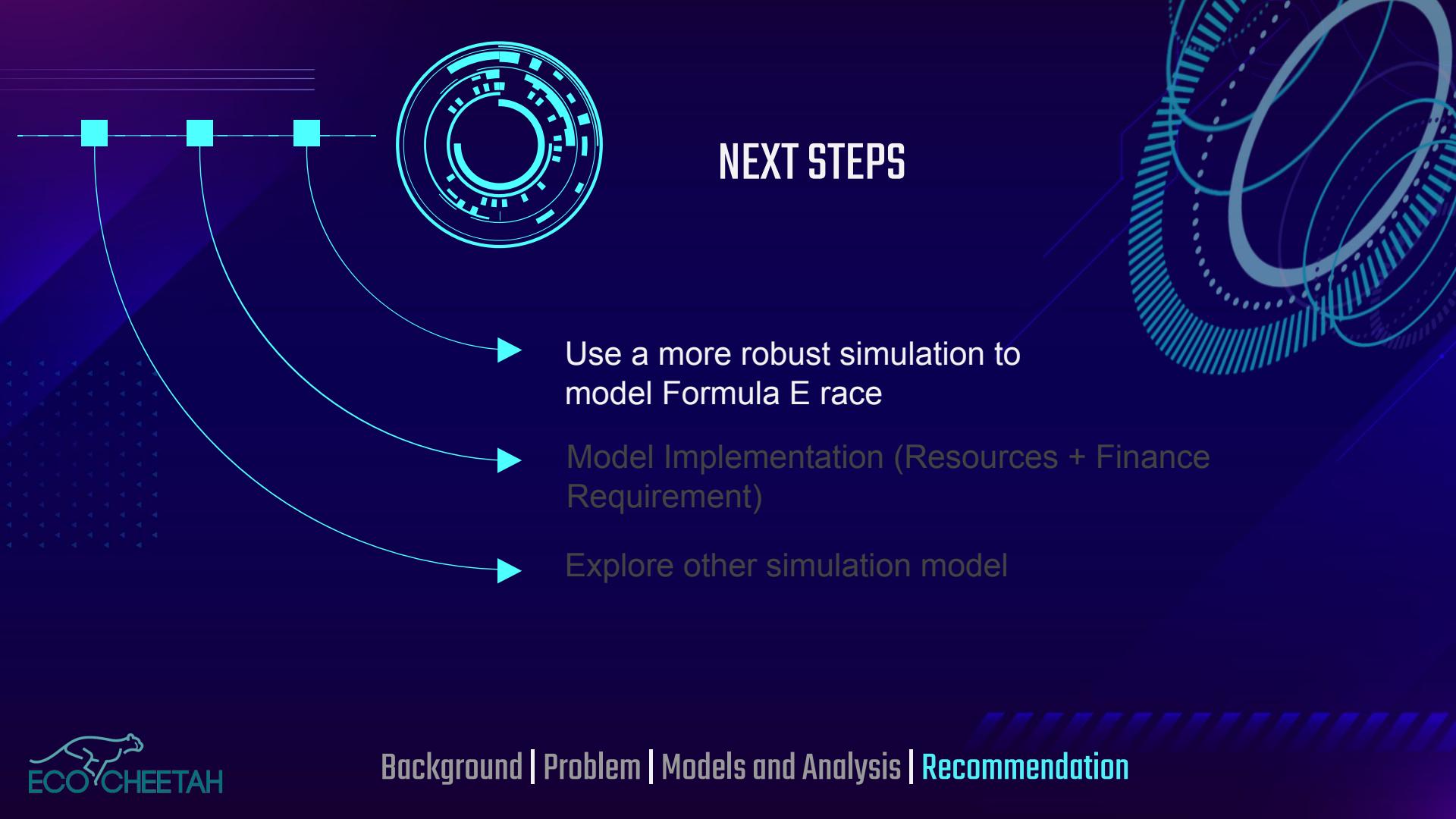
WEAKNESS

Uncertain key factors in
simulations

OPPORTUNITIES

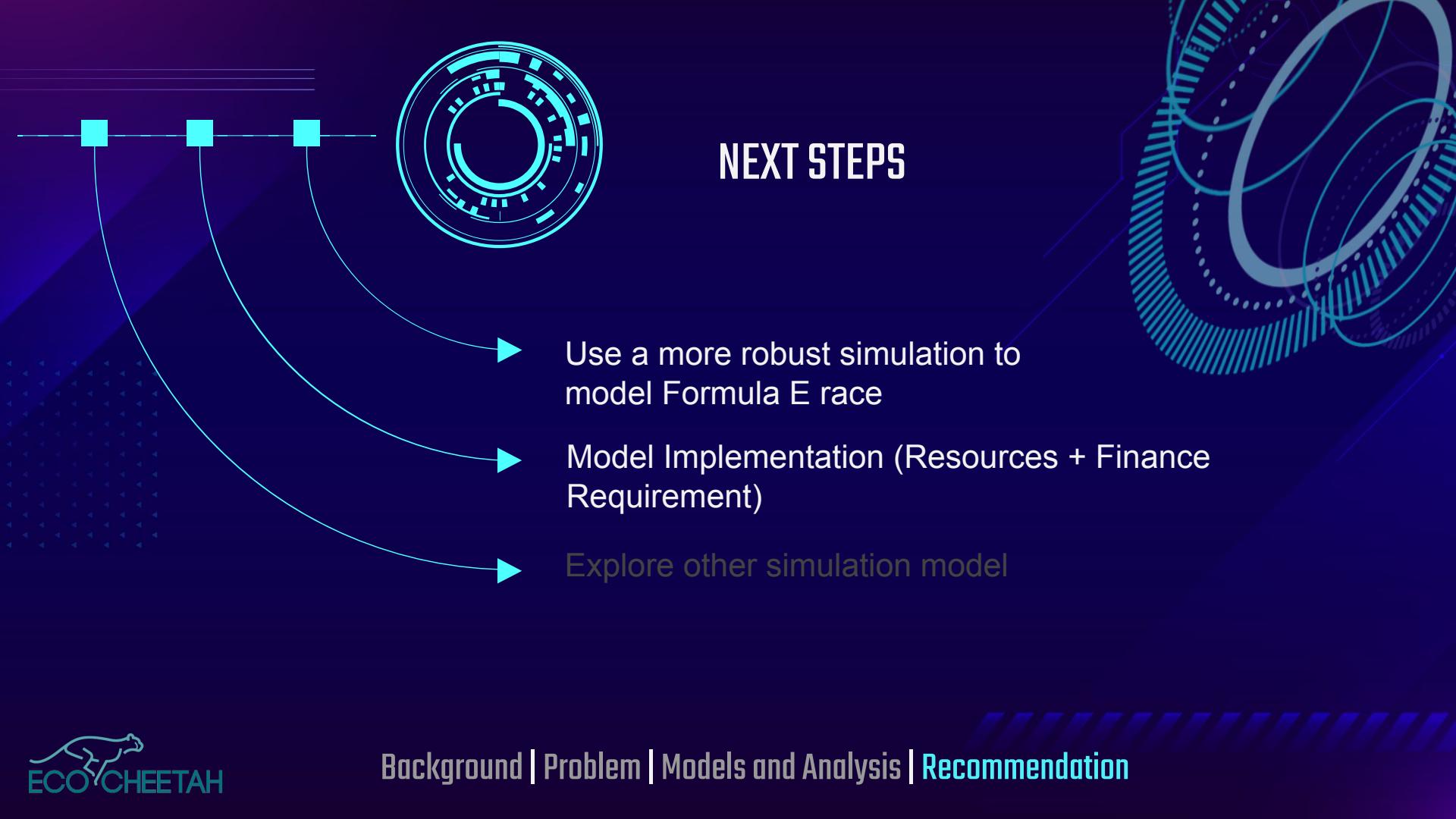
Data + Analytics = Impact





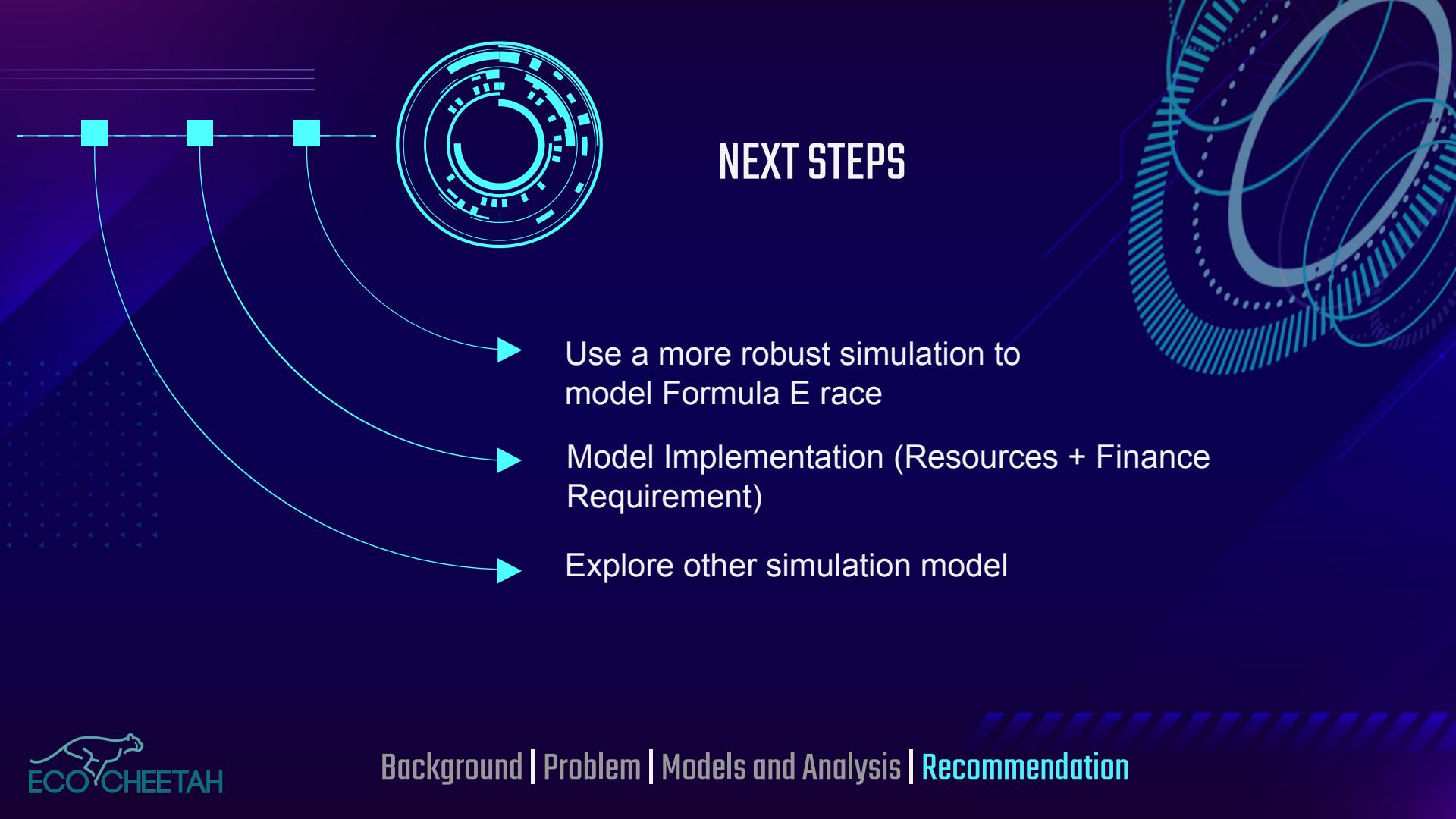
NEXT STEPS

- ▶ Use a more robust simulation to model Formula E race
- ▶ Model Implementation (Resources + Finance Requirement)
- ▶ Explore other simulation model



NEXT STEPS

- ▶ Use a more robust simulation to model Formula E race
- ▶ Model Implementation (Resources + Finance Requirement)
- ▶ Explore other simulation model



NEXT STEPS

- ▶ Use a more robust simulation to model Formula E race
- ▶ Model Implementation (Resources + Finance Requirement)
- ▶ Explore other simulation model



ANY QUESTIONS



SOURCE CITED

- ◀ Liu, X., & Fotouhi, A. (2020). Formula-E race strategy development using artificial neural networks and Monte Carlo tree search. *Neural Computing and Applications*, 1-17.
- ◀ Pitch Deck Presentation templates. (n.d.). Retrieved November 12, 2020, from <https://slidesgo.com/pitch-deck>
- ◀ Robinson, S. (n.d.). Decision Trees in Python with Scikit-Learn. Retrieved November 12, 2020, from <https://stackabuse.com/decision-trees-in-python-with-scikit-learn/>



ECO CHEETAH

THANK YOU

11/13/2020