

How Data Science Research Can Serve the Region

Toolbox Training – 5/12/2020

Lyle Janicek, SCAG

Jeanne Holm, City of Los Angeles

Anthony Lyons, City of Los Angeles



www.scag.ca.gov



Agenda

1. What is Data Science
2. Introduction of Data Science Federation
3. Student Presentation – Los Angeles Zoo Project
4. Data Science Federation 2020/2021 Call for Projects
5. Q/A Session

What is Data Science?

Data science is the study of data. It involves developing methods of recording, storing, and analyzing data to effectively extract useful information. The goal of data science is to gain insights and knowledge from any type of data —both structured and unstructured.

How and why do we use it in Planning?

Now, more than ever, planners are relying on data to help determine policies and decision making.

Data Science Case Study- City of Boston

The city has embraced data-driven management with a program called CityScore.

The initiative consists of an online dashboard that shows how the government is performing relative to its goals in 24 key areas, such as responding to emergency calls and collecting garbage.

Sensors automatically record much of the data, and city workers record information on mobile apps when they complete an activity. The system makes problems apparent and helps ensure the city allocates resources where they will have the most impact for citizens.

Similarly, The City of Los Angeles has created the LA Data Science Federation to provide data related resources for city staff and the community.

City of Los Angeles Data Science Federation

Jeanne Holm, Chief Data Officer, City of Los Angeles

Introducing the City of LA Data Science Team

The Data Science Federation

- Started in 2016 to better support students in data science with applied city challenges, to recruit data scientists to government service, to educate city leaders on new data science techniques and ideas, and to solve difficult city challenges

Data science activities in the City of LA

- Data and Predictive Analytics Group
- Data Science Federation
- Data Angels
- Data Academy



Our DSF Community- Join Us!



Cities and Governments

- City departments
- Elected officials
- County government
- Regional groups

Educational Partners

- Universities
- Colleges
- Community groups
- Boot camps
- ...and more!

Scientists and Hackers

- Coder and hacker groups
- Volunteers
- Interns
- Employees

By the Numbers

18 universities and colleges

54 partners

88 cities

40+ projects

300+ students engaged



Submitted by [Higher Education](#) on March 10, 2017
Art Center College of Design



Claremont Graduate University



Cal State Long Beach



Pepperdine University



Occidental College



Cal Poly Pomona



Cal State Dominguez Hills



Cal State Northridge



University of California Los Angeles



Loyola Marymount University



California Institute of Technology



Cal State Los Angeles



Los Angeles Trade Tech

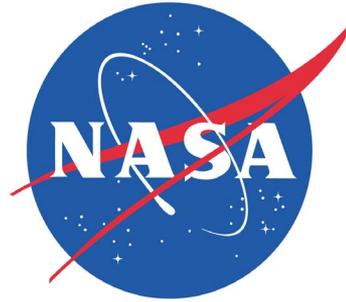


University of Southern California



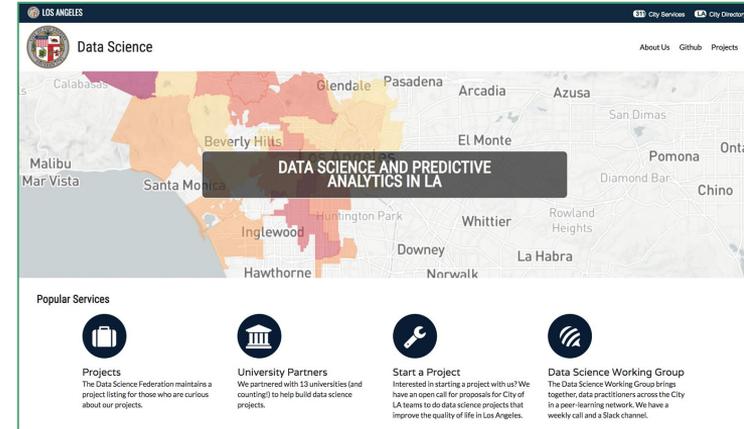
How Universities and Colleges Participate

1. Express interest
2. Some host a hackathon or data day
3. Create a data sharing agreement
4. Respond to project calls
5. Circulate job postings
6. Collaborate on grants
7. Submit for regional and global awards



What is a DSF Project?

1. Can be completed in less than a year, often within one quarter or semester
2. Does not involve very sensitive data
3. May graduate into a larger project, grant proposal, or job opportunities for students
4. A City + Agency Project Manager + University Team = One project
5. Generally proposed by City departments, but open to university generated ideas



Evolution of the Federation

In 2018

- Expanded from 188 cities and added three new universities

In 2019

- Get more cities actively working on projects
- Staff up City of LA team to bring new skills and operationalize the process

In 2020

- Expand the model for local governments to build federations and share code, projects, and talent
- Integrate emergent issues for smart cities



Zoo Business Research

Sonya Zhang, PhD., CGU
Yang Zhong, PhD Student



Los Angeles Zoo

School Visitation Analysis

Abdulmajeed Mesefir
Chien-Ting Chang (Ted)
Yang Zhong



Agenda

- Objectives
 - 10 Years Visitation Data Analysis
 - Trends & Patterns
 - LA County Schools Data Analysis
 - Recommendations



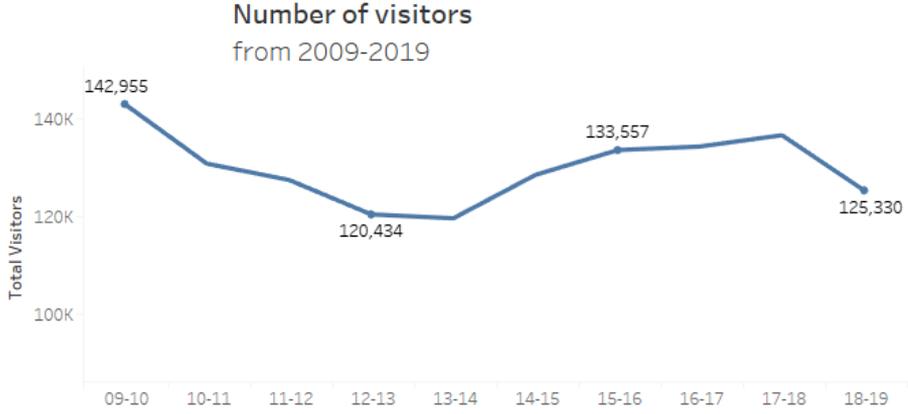
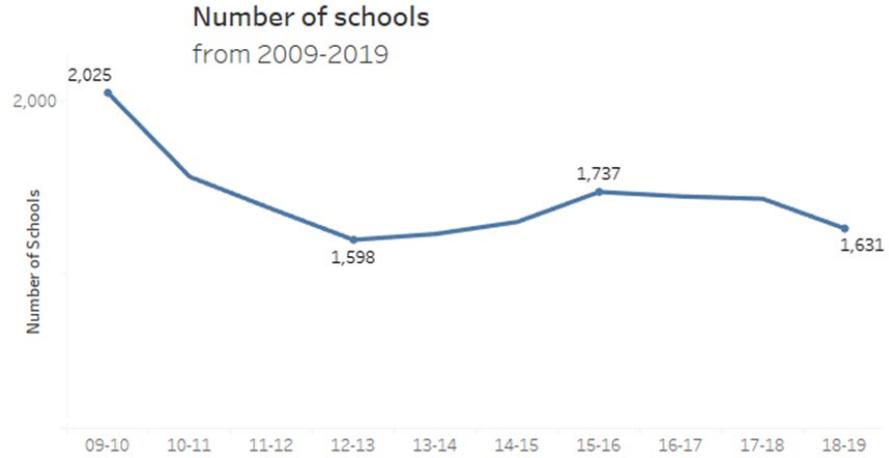
Objectives

- School visitation in past 10 years
 1. Overall trends & visitor composition
 2. Geographic patterns
 3. Significant factors that influence visitation
- Which schools are not visiting the Zoo
 1. The schools which never came
 2. The schools which came before but stop visiting in recent years



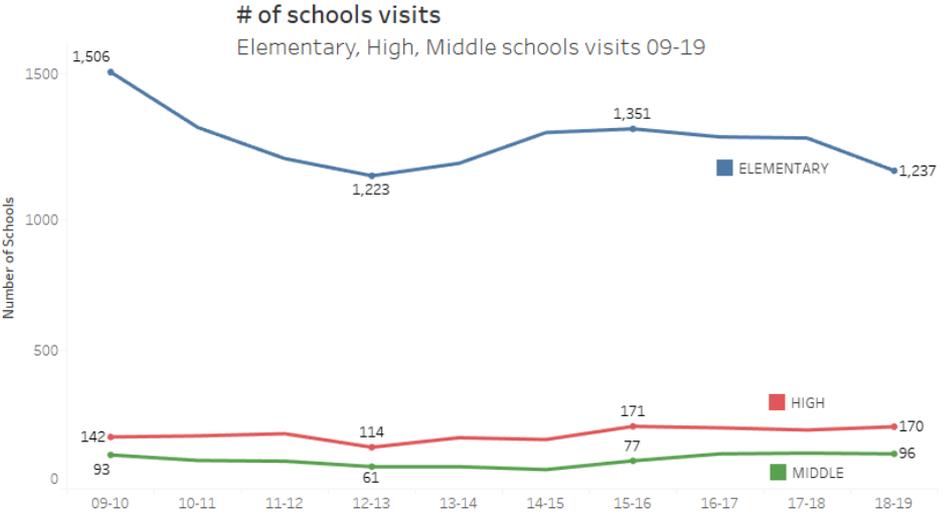
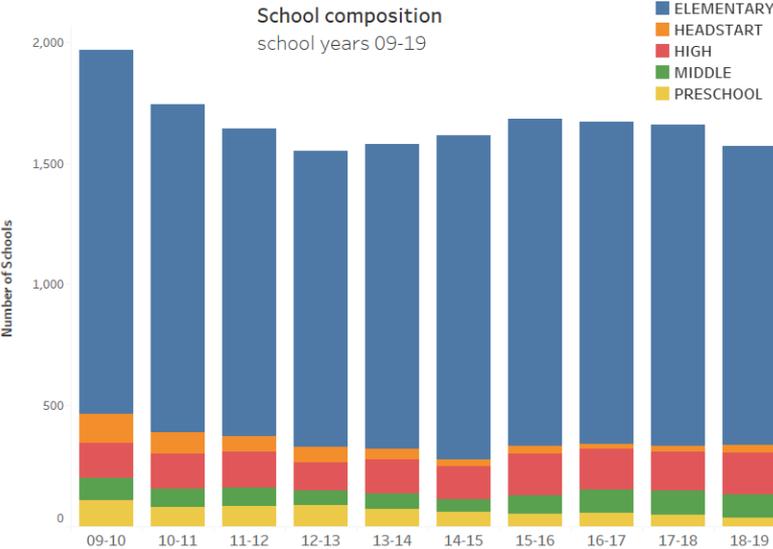
The Overall Trend

- The best year is 09-10
- Then decreased a lot until 12-13
- Recover from 12-13 to 15-16
- Been descending in recent years



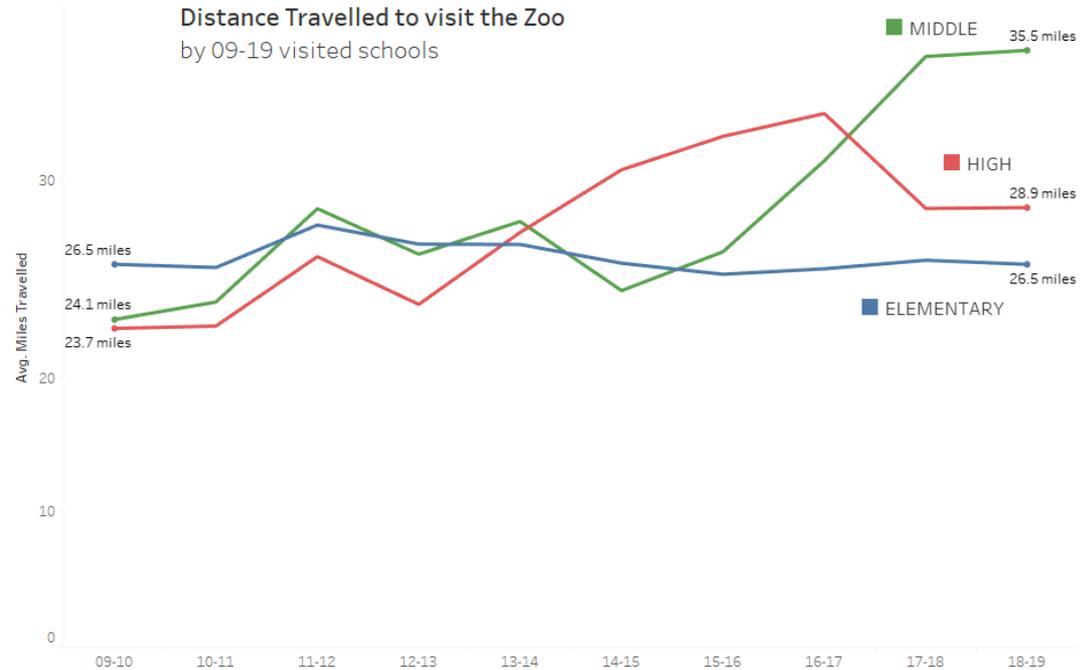
The Visitor Composition

- Elementary schools contribute 80% of the visits; (2nd - High schools, 3rd - Middle school)
- Elementary school visitation changes



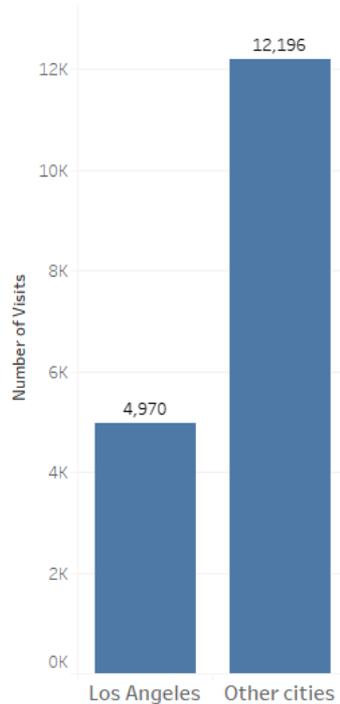
The Geographic Pattern

- Elementary schools distance has been consistent
- Middle & High schools from further distance started visiting in recent years
- In the first recession period, the closer schools stop visiting



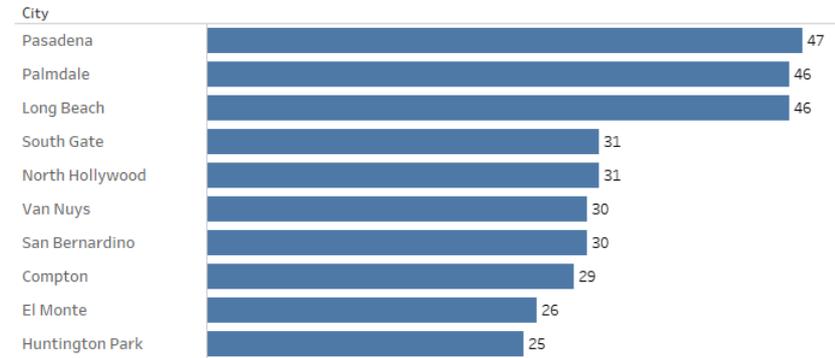
- Schools from the City of LA contributed about 30%

Visits from LA or non-LA cities
09-19

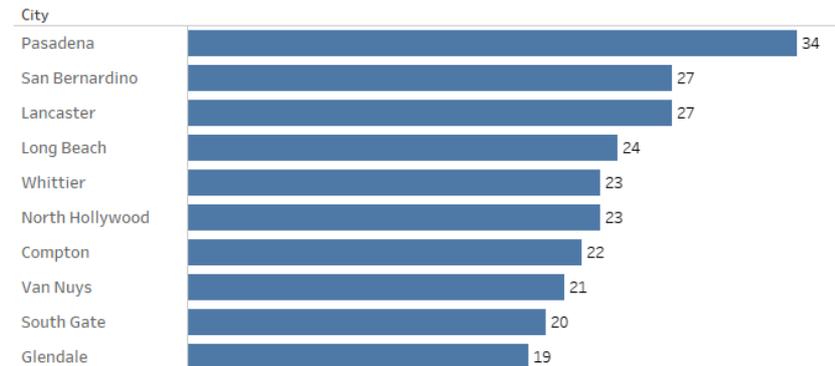


- Visitations distribute evenly than before

The top 10 cities (Except LA) that visited the Zoo from 09-10



The top 10 cities (Except LA) that visited the Zoo from 18-19

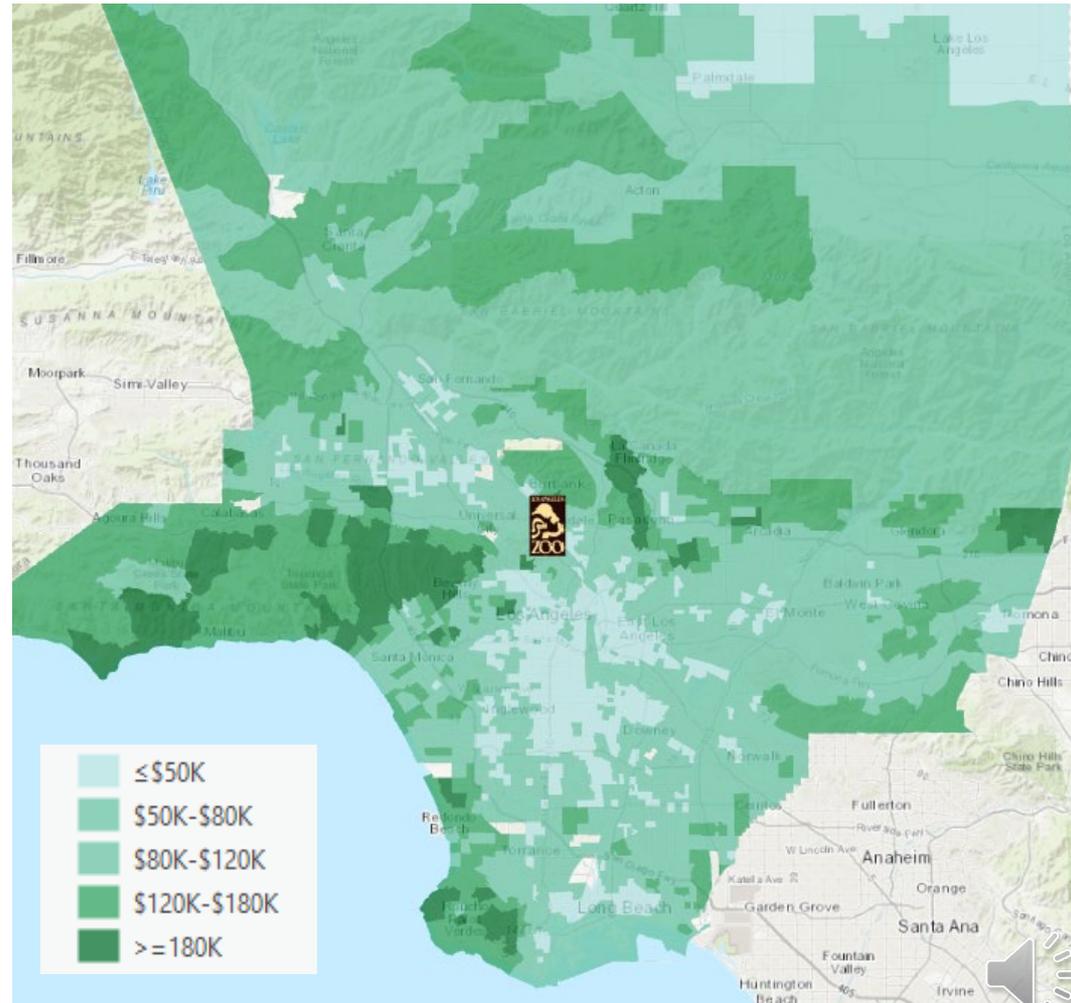


The Significant Factors

Demographics - Income

- Median Household income within 5 miles of the school
- The schools located in lower income area seems more likely to visit

	Visited	Total	%
<50K	612	857	71.41%
50K-80K	650	1019	63.79%
80K-120K	505	829	60.92%
120K-180K	195	361	54.02%
>180K	37	81	45.68%

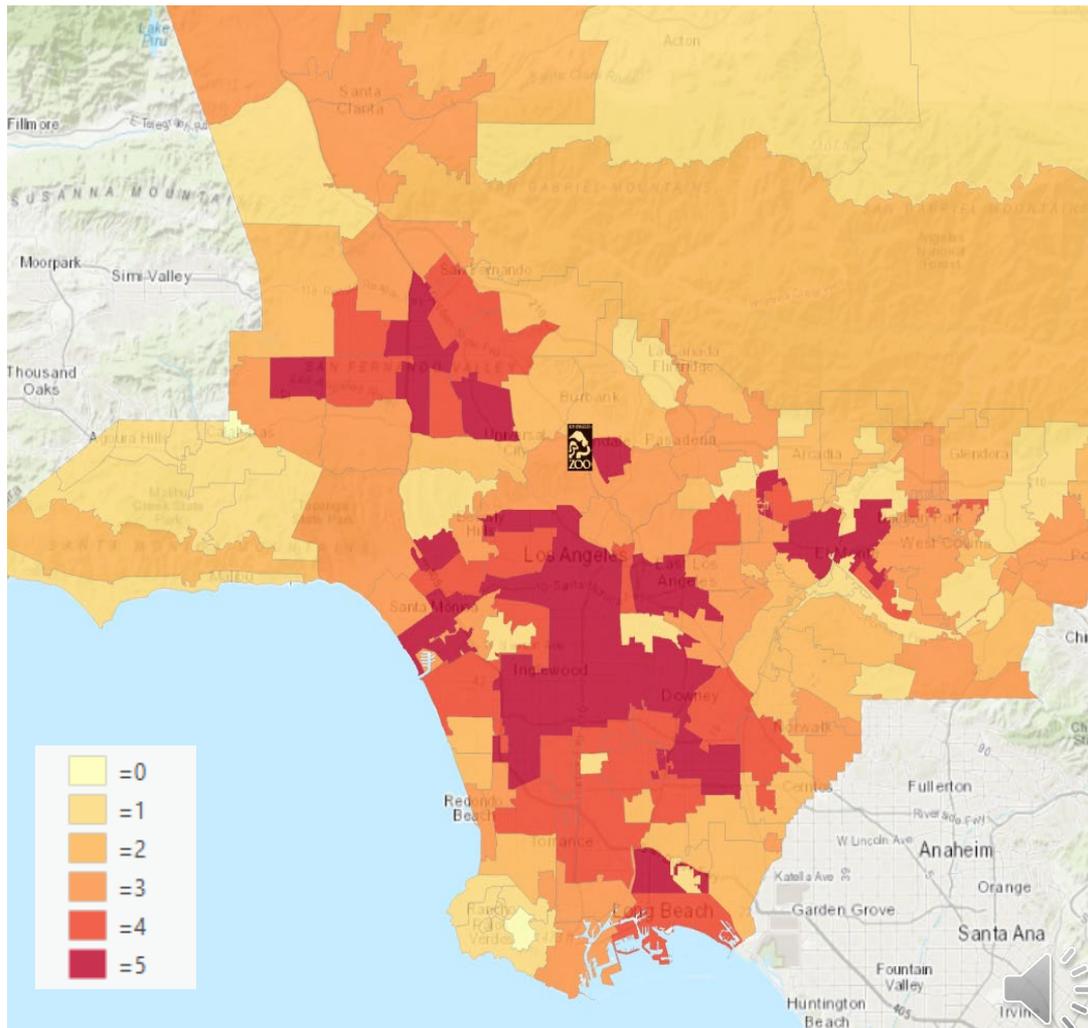


The Significant Factors

Park Need Level

- A score that represents how much the area need parks
- The score ranges from 0 to 5 (low to high)
- The schools located in high park need level area did visit more

Park Need Level	Total School	Visited	%
1	191	86	45.03%
2	548	318	58.03%
3	833	510	61.22%
4	591	393	66.50%
5	981	690	70.34%



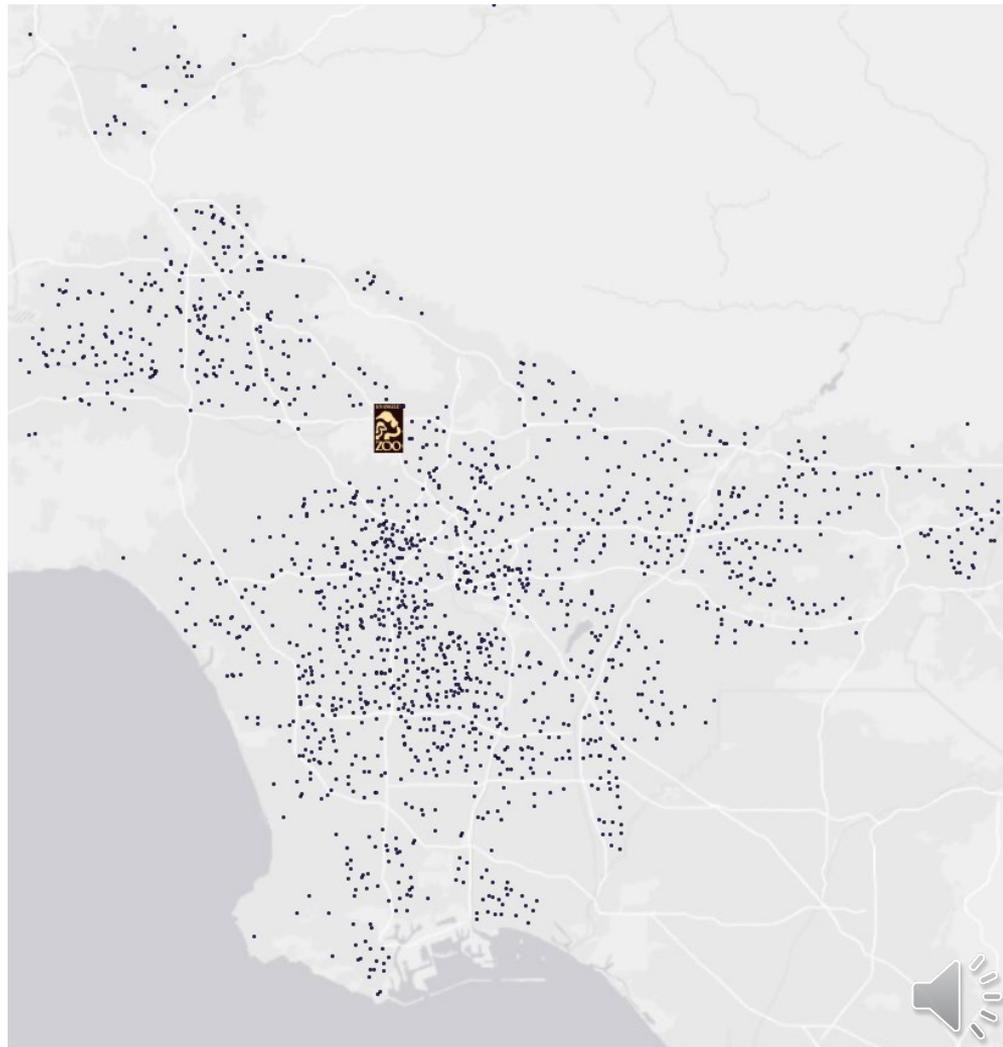
Other Factors - The Title 1 Status

- The schools with large number of low-income students can receive federal funding
- 1,780 out of the 3,147 schools in LA country are 'Title 1' (56.5%)
- The title 1 schools are more likely visiting the Zoo than the non-title 1 schools

	Visited	Total	%
Title 1	1277	1780	71.74%
Non-Title 1	722	1367	52.82%

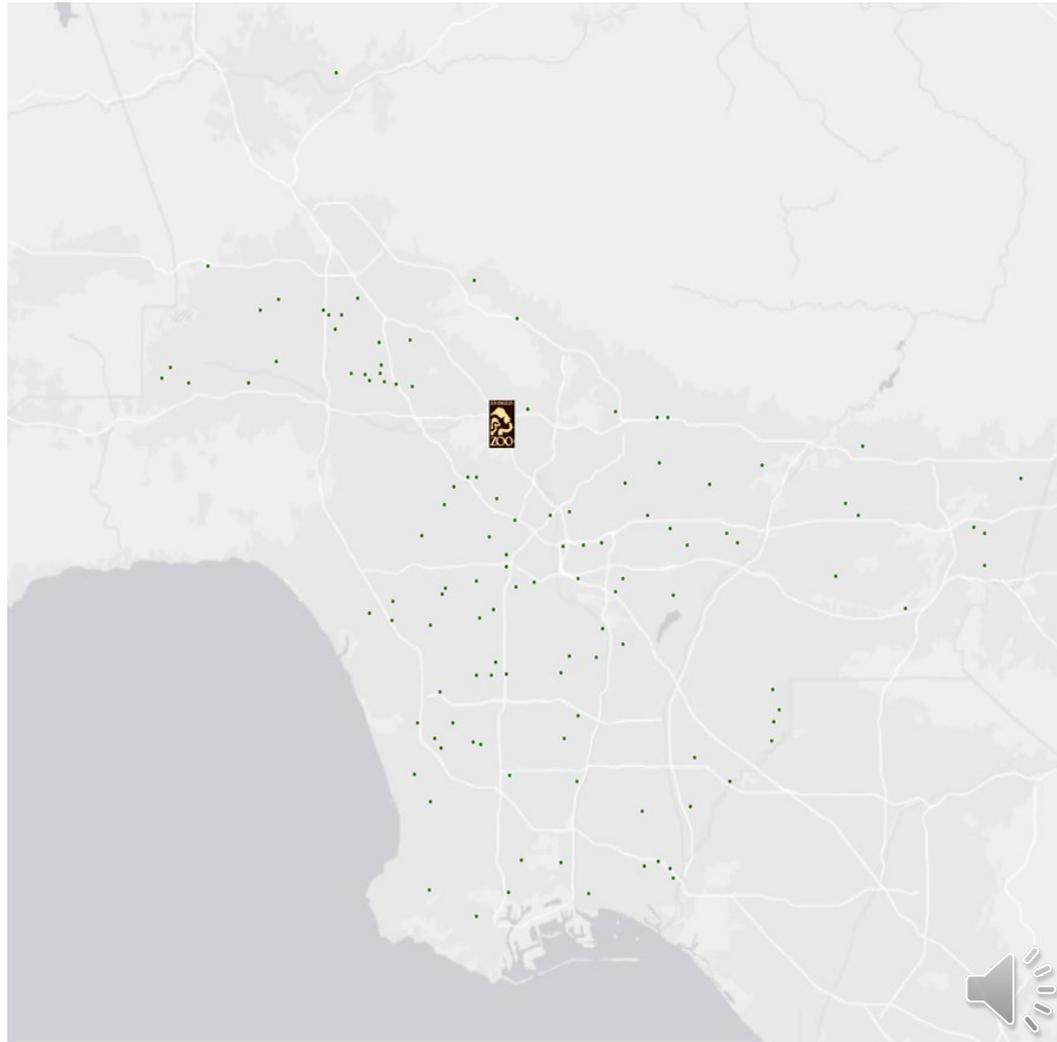
Last 3 Year Visited		Total	%
Title 1	923	1780	51.85%
Non-Title 1	479	1367	35.04%

Avg. Visited/School	
Title 1	6
Non-Title 1	4



Schools not visiting the Zoo

- 3,147 schools in LA county (CDE Directory)
- 2,125 of them are Elementary schools
- Schools never visited the Zoo are **586**
- 1,539 schools have visited at least once.
- In the visited schools, **127** of them used to visit the Zoo in the past (>3 times) but not visiting now (within the last 3 years)

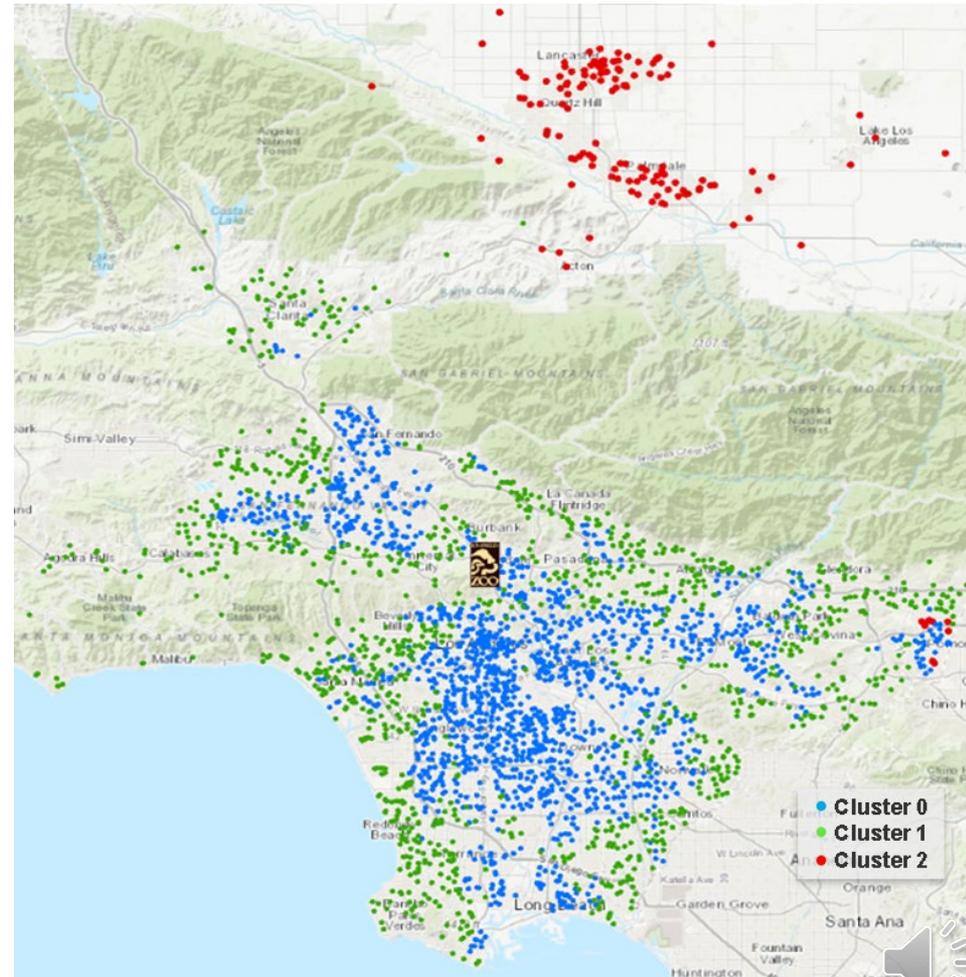


School clustering

Applied data clustering to 3,147 schools by

- Distance to the Zoo
- School rating
- Surrounding median household income
- Park need level
- How many times visited the Zoo (10 years)

3 clusters were formed



Recommendation - characteristics of the schools (more likely to visit)

Cluster #0 has the best record of visiting the Zoo

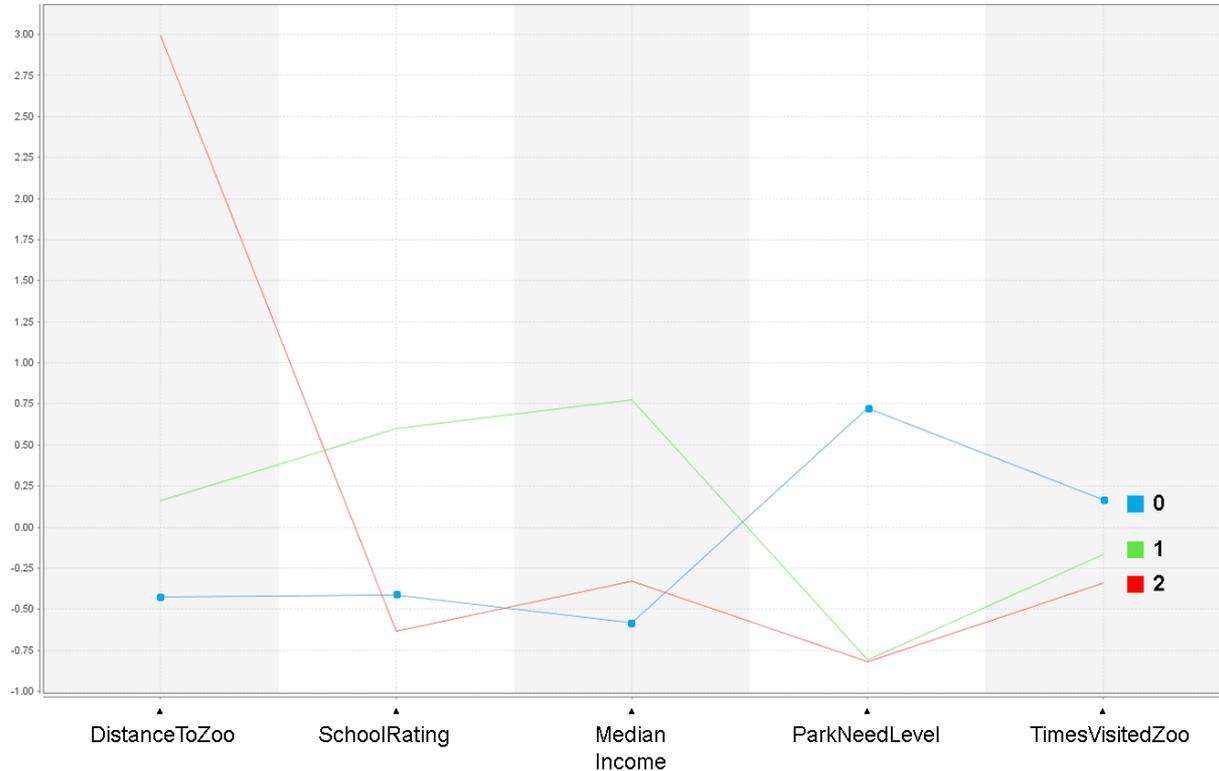
Cluster #0 characteristics are:

- Close to the Zoo
- Average on school rating
- In low income areas
- In high park need level areas

1,759 schools in cluster #0

550 of them never came

108 of them used to visit before but not in recent 3 years



Conclusion

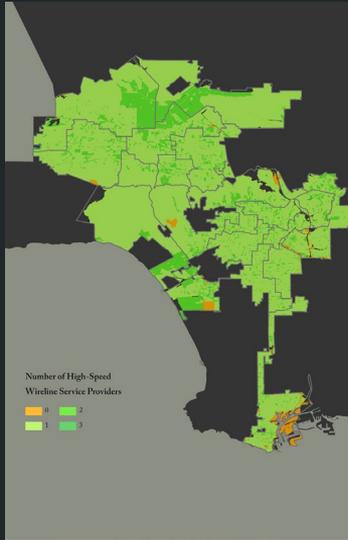
- Target the **elementary schools** which never came and the ones came before but stop visiting in recent years (a total 713 schools)
- Target the schools which never came and the ones came before but stop visiting in recent years (a total 658 schools) in **cluster 0**
 - There are 327 schools which appear in both list





Thank you





DSF Call for Projects 2020-2021

City of Los Angeles / Southern California Association of
Governments

Project Status for 2019

25 projects submitted

17 that met all requirements

12 matched with Universities

2 projects handled internally

Project management was handled by sponsoring organization or city, not by the City of L.A.'s ITA or SCAG

Goal: Connect Civic Data Science Needs to University Resources

Revised Onboarding Process

New Call for Projects for Academic Year 2020-2021

Targeting 1 year project duration

Shared Pipeline - Individual Scoping Meetings

Project Scoping

- 1) **About** - Short description of the project
- 2) **Lead Department/Agency** - Who will lead the project
- 3) **Partnering Departments** - Any other departments or agencies involved
- 4) **Team Members** - Description of staff that will be engaged on the project
- 5) **Goals** - How project will solve a problem
- 6) **Deliverables** - Tangible outcomes (e.g. code, application, report, etc.)
- 7) **Data** - What data will be used (link please!)
- 8) **Timeline/Sprints** - How much time will the project take

Goals vs Deliverables

Goals are:

- Examine data to understand hotspots for pedestrian collisions and compare to built environment to determine causes.
- Develop code to analyze large data sets related to homelessness to determine social service needs based on geography.
- Understand the lack of park access and transportation options for reaching regional open spaces for a low income community.

Deliverables are:

- A report explaining findings
- A cleaned data set and code for analysis
- An application that uses data to perform a function (e.g. dashboard)

University Onboarding Process

University Outreach - add relevant staff/professors to DSF listserv

University teams “bid” on a project of interest by providing resources - how many students, how long, etc

NDA or Data Sharing Agreement signed with partners

(Optional) Joint Grant Applications

Deadlines

Early May: Call for Projects opens

May → June: Project scoping sessions, submit data for review

July 12: Initial scopes due

July 31: Final review of projects

Timeline

August 5: Projects sent to universities

August 30: Initial university responses due

Mid-September: Project charters and project matching finalized, initial meetings setup

Ongoing: Projects start

Who is this process for?

Process is for Medium sized projects

Targeting 1 year (should be completed within) and mostly shooting to have students either doing Masters Thesis, Senior Colloquiums, Internships. Etc

If large project, needs either additional support or grants.

FAQ

Q: Does this cost any money?

A: The Program is free to participate in, however, you are required to designate and commit staff time to the project.

Q: Who does the matching?

A: ITA and SCAG Staff based on the needs you identify in the scope document.

Q: What happens to the project when the students leave?

A: All projects culminate in a final project presentation and handoff, and all code should be documented and built on Github for reproduction and enhancement

Q: Do you support grant funding?

A: Yes, DSF projects have received \$500K + worth of grant funding from multiple sources

Q: Do you work with Personally Identifiable Information?

A: Yes but that typically requires a unique data sharing agreement and scope process.

Q: How many projects have been completed in the past?

A: Approximately 25

Questions/Comments?

Contacts

SCAG

Lyle Janicek, Associate Regional Planner

janicek@scag.ca.gov

City of Los Angeles

Jeanne Holm, Chief Data Officer

jeanne.holm@lacity.org

datascience@lacity.org

Links

Website : dsf.lacity.org

[DSF Factsheet / FAQ](#)



Thank you!

Upcoming Toolbox Training:
California Healthy Places Index Tool
Monday, May 18, 2020
10:00 a.m. – 11:00 a.m.

Register here:

<https://scag.wufoo.com/forms/zk9m3x61949avd/>