WHY DO PEOPLE GO TO PARKS?

Exploring the factors that drive park attendance

Author: Angela Zhao, Data Scientist

City of San José, Information Technology Department

May 2, 2024

Acknowledgements

This project was made possible by collaborative effort between multiple City of San José departments. The individuals below were key to making this possible.

Claire Cioni, Program Manager, PRNS

Lisa Diebolt, Analyst II, PRNS

Albert Gehami, City Privacy Officer, ITD

Arti Tangri, Equity Data Lead, ITD

Nolan Bertuca, Senior Geographic Information System Specialist, PRNS

Justin Imamura, Senior Events Coordinator, PRNS

Brian Clampitt, Parks Manager, PRNS

Arlene Gonzalez, Analyst I, PRNS

Alvaro Celaya, Events Coordinator II, PRNS

Anthony Morales, Events Coordinator II, PRNS

Chris Boyer, Recreation Program Specialist, PRNS

Denise Esparza, Senior Recreational Leader, PRNS

Avi Yotam, Deputy Director, PRNS

Khaled Tawfik, CTO, ITD

Mitsuki Irihara, Data Analyst Intern, ITD

Table of Contents

Infographic:

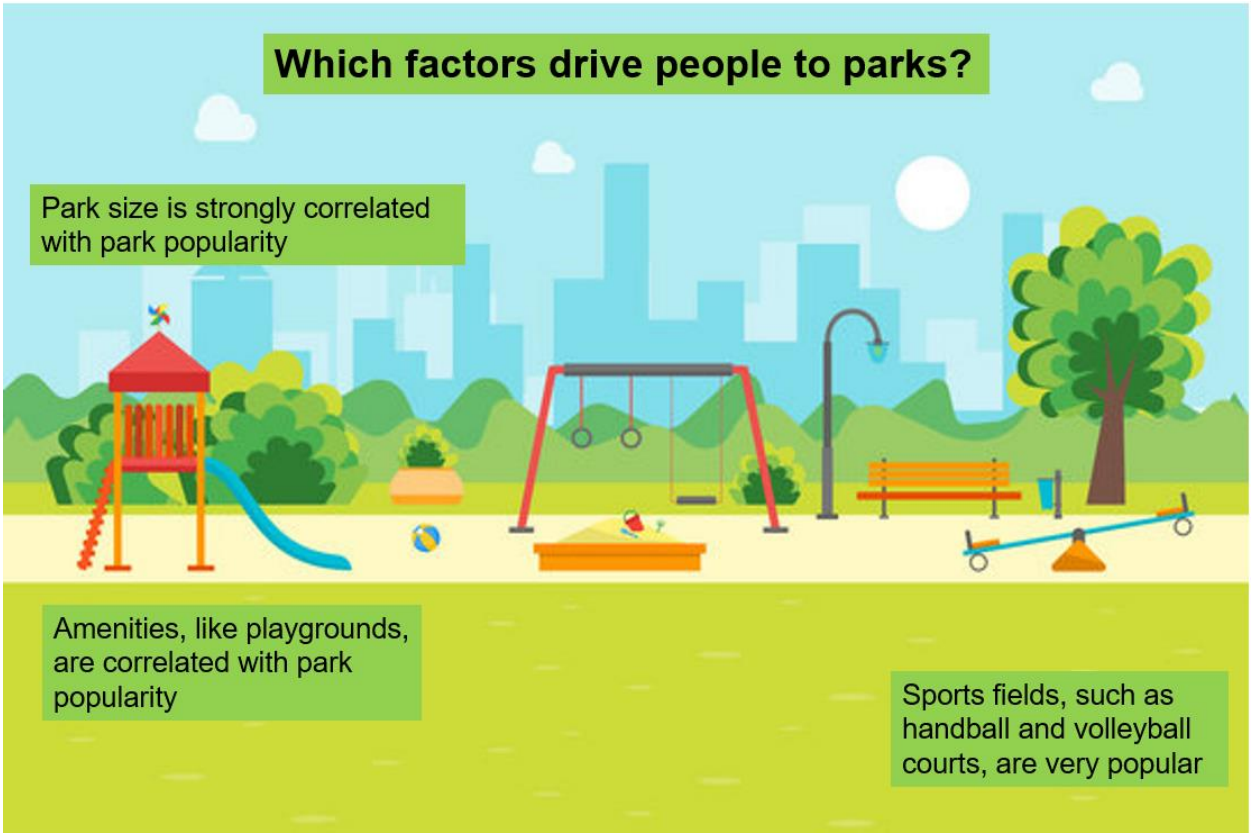

Figure 1: Source: https://www.123rf.com/clipart-vector/park.html

Overview:

Higher park use is correlated with better health[1]. By understanding which key factors drive people to parks, the Parks, Recreation, and Neighborhood Services (PRNS) department can improve park quality by better strategically allocating resources for park development and maintenance. With this goal, PRNS set out to **identify which factors drive people to parks, and of these factors, which are most deserving of investment.**

The investigation found that group activities are one of the key drivers of park attendance. Amenities that enable group activities, such as sports fields which allowed for pick-up sports games, are strongly correlated with increased park attendance. Therefore, the recommendation is to increase investment into maintaining and building amenities. These include non-reservable sports fields, such as handball courts, and games.

There are 201 parks in this analysis of foot traffic data for year 2022, but only 1/3 of the parks have sports fields. The analysis looks at visitors from the 227 census tracts and 634 census block groups in the City. Parks were either regional, neighborhood, or dog parks.

[1] Grilli, G., Mohan, G., & Curtis, J. (2020). Public park attributes, park visits, and associated health status. *Landscape and urban planning*, *199*, 103814.

Background

Previous Work done by the City of San José

The City of San José is interested in learning the factors that drive park visitation to build and maintain parks that better serve the residents.

A previous analysis was conducted on the factors driving park attendance within a 10-minute walk of residence with the Stanford Future Bay Initiative. The analysis identified 32 factors that accounted for 36% of the variance of a park's average daily visits[2]. However, the size of the park accounted for 17% of the variance, which leaves 31 demographic factors to account for the remaining 19%. The analysis does not suggest that there is a single demographic factor that is strongly correlated with park visits.

Additionally, previous work has also included descriptive analytics of park visits. The analysis uses the same foot traffic data as this analysis. Findings concluded that foot traffic peaked at parks from 9-10am and 5-6pm – the start of the workday and after work. Peak days are Sunday, Saturday, and Wednesday. Finally, the most popular time of year is from mid-June to mid-July. As a last finding, the analysis concluded that there is a slight negative relationship between the total time spent at the park and how well maintained the park is. This odd relationship suggests that popular parks are so popular that maintenance is not keeping up with demand (i.e., the amount of wear-and-tear).

Previous work also tested if demographics had a relationship with park visits in relation to resident backyard size. The City will continue testing demographics. While it is unlikely that a single demographic factor is strongly correlated with park visits, it is possible that a variety of factors are.

Background Literature

Background literature suggests that the strongest correlates of park visitation are park size and the number of organized activities observed, such as sports games or picnics. Park amenities (such as sports fields and playgrounds), Amenities, accessibility, and safety were also considered. A research paper by Cohen et al 2009 compared parks between those with improved and new amenities and those without, concluding that an estimated "39% of decline in park users could be directly attributed to a reduced number of organized programs" [3]. To underscore the importance of organized activities, a park with low safety ratings and a "shooting/attempted murder incident" experienced increases in park visits due to "the scheduling of additional baseball games".

[2] Wagenfehr, J. (2020, August 10). San Jose Park Accessibility Analysis. stanfordfuturebay.github.io. https://stanfordfuturebay.github.io/sj_park_accessibility_analysis.html

[3] Cohen, D. A., Golinelli, D., Williamson, S., Sehgal, A., Marsh, T., & McKenzie, T. L. (2009). Effects of park improvements on park use and physical activity: policy and programming implications. *American journal of preventive medicine*, *37*(6), 475-480.

Another paper by Cohen et al 2009 also concluded that factors not correlated with park usage include "neighborhood population density, neighborhood poverty levels, perceptions of park safety, and the presence of a park advisory board"[4].

Background literature suggests that organized activities and park size contributed the most to park foot traffic. In combination with the City's previous work on park foot traffic, **the analysis proceeded to explore several factors: resident demographics, organized activities, and park amenities.**

Analysis:

Data

Several data sources were used for the analysis. The analysis was limited to only the parks in the Healthy Places Index and only the park visitors living in San José. Sources used in the analysis are listed here:

1. The Anonymized Carrier Data Movement dataset containing weekly foot traffic information for various areas in the City, including parks and includes data for visitors' census block groups (CBG). The foot traffic data is collected by tracking mobile devices and is used as a proxy for park visitation. The vendor has previously cleaned the dataset to reduce noise. In the analysis, the data is further cleaned by removing all visits of duration less than 5 minutes. Parks near freeways or busy streets demonstrated higher park visits lasting less than 5 minutes, possibly due to cell phone data capturing people driving or passing by the park instead of visiting the park.
2. The Healthy Places Index was used to determine which parks would be analyzed – parks that had a Healthy Places Index were included in the analysis. The Healthy places Index maps data on social conditions that drive health.
3. Most recent census data pulled from the 2021 5-year American Community Survey. Income level, race, education level, age, percentage of households with children, and gender was used to answer the question if different populations were using different kinds of parks (categorized by popularity).
4. The schedule of park events, sports reservations, and picnic reservations were used to determine a subset of organized activities, which had been a potential factor driving park usage.
5. The amenities at each park were also categorized and analyzed for their relationship to foot traffic. This helped determine which amenities were most correlated with high park visitation and which amenities should be maintained more regularly or added to more parks.

[4] Cohen, D. A., Marsh, T., Williamson, S., Derose, K. P., Martinez, H., Setodji, C., & McKenzie, T. L. (2010). Parks and physical activity: why are some parks used more than others?. *Preventive medicine*, *50*, S9-S12.

Park Categorization:

The analysis began by separating parks into four categories: broadly popular, locally popular, less popular, and moderately popular parks, since these parks have different visitation patterns.

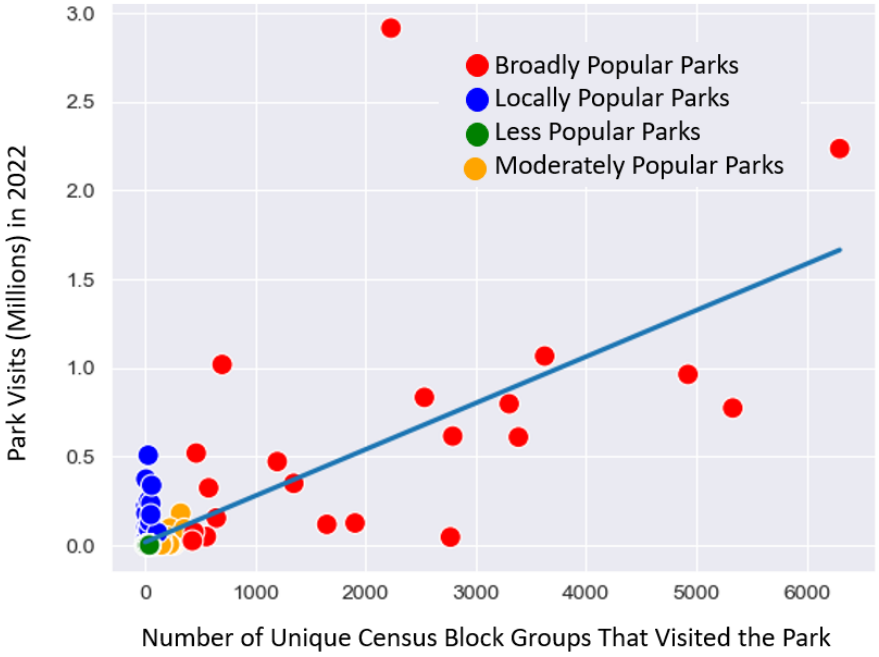

Figure 2: Park popularity categorization based on how many CBGs visited the park vs Park visits

The number of unique census block groups (treated as neighborhoods) that visited each park was used as a proxy for park popularity – essentially, how far away are people coming from to get to the park? For reference, there are 634 census block groups (CBGs) in San José and 227 census tracts. The analysis included visitors from outside of San José when looking at park popularity. There were about 7000 unique CBGs recorded in the data.

1. **Broadly popular parks (red)**: 22 parks had more than 400 CBGs visit the park in a year. Most of these parks received more than 100% of expected visitors based on the number of CBGs that visited the park. Due to the large number of CBGs that visited the park, the expected multiplier of visits was relaxed.
2. **Locally popular parks (blue)**: 22 parks that received at least 90% of expected visitors based on the number of CBGs that visited the park and had less than 150 CBGs visit the park in a year.
3. **Less popular parks (green)**: 26 parks that received less than 10% of expected visitors based on the number of CBGs that visited the park and had less than 50 CBGs visit the park in a year.
4. **Moderately popular parks (yellow):** 131 parks that performed as expected and did not fall into any of the prior groups.

All the specific parks in each category can be found in the appendix.

Each type of park also has significant differences in park usage, including average dwell time, distances traveled to the park, and average number of repeat visits to the park. The Welch t-test and the student t-test were used for significance testing[5].

Demographic Drivers of Park Attendance

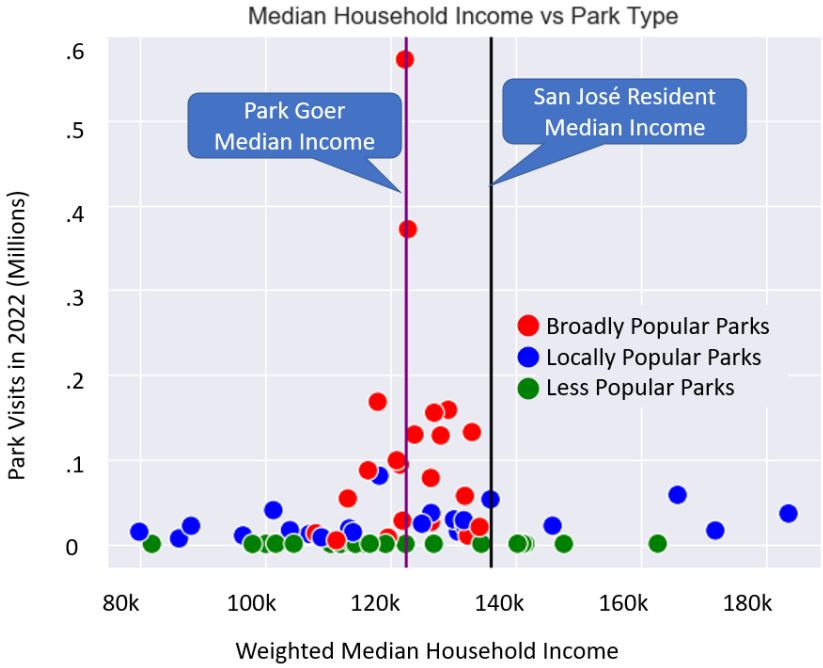

Figure 3: Park goers generally make less income than the average San José resident.

The strongest factor correlated to park visitors is park size and income. The additional factors tested, demographics and population density, had little correlation to the number of visits around the park. This indicates that more people living near a park does not necessarily mean more people will go to that specific park, and that park popularity is not due to a trait that is strongly correlated with a specific demographic.

Population density was calculated by taking the population density of the census tract the park was located. Population density has little correlation to park visits, suggesting that how close a park was to a resident did not factor strongly into how likely the resident was going to go to the park. This may also indicate that certain parks have attractive features, such as amenities, that draw visitors.

The relationship between resident demographics and park categories was explored to see if there were qualities in different parks that appealed to different demographics. The factors tested included: household income, race, households with kids, and age. Chi-squared tests were used to determine significance[6]. There was no significant correlation between different types of parks and any of the demographic factors, indicating factors that determine park popularity are not built on factors that have demographic appeal.

The analysis suggests a relationship between income and park attendance. Park goers were more likely to earn less than the median income of San José residents. Furthermore, broadly popular parks and less popular parks goers make less than the median income of park goers in general. This indicates that parks may offer low-cost activities, such as sports or gym equipment, that can otherwise be difficult to access at lower incomes.

[5] West, R. M. (2021). Best practice in statistics: Use the Welch t-test when testing the difference between two groups. *Annals of clinical biochemistry*, *58*(4), 267-269.
[6] https://www.investopedia.com/terms/c/chi-square-statistic.asp

The City hypothesizes that park goers are lower-income residents.

Organized activities

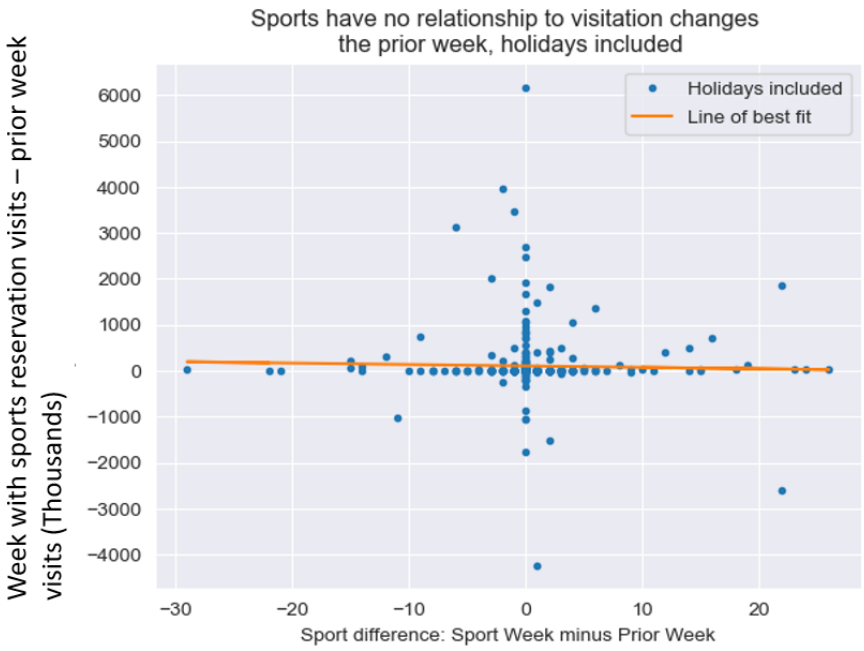

Figure 4: Sports reservations have no relationship to changes in visitation.

Organized activities covered three categories: events, sports, and picnics. Events are planned by the Placemaking team under VivaParks, which aims to promote community engagement in local neighborhood parks by hosting activities such as painting or lawn games. Sports and picnic fields can be reserved by individuals or organizations, usually for activities such as sports competitions or birthday parties.

Not all parks have reservable sports fields and picnic sites. Additionally, reservable sports fields only cover softball, soccer, tennis, and baseball fields at 42 parks, which is only a subset of sports fields in about a quarter of the parks. These sports fields are aimed at sports leagues that practice and hold competitions, which is not a neat reflection of neighborhood usage of parks. This caveat carries over to picnics as well, since only a subset of parks has reservable picnic sites.

As a result, the organized activities section can be better discussed as the effect of activities organized by larger organizations. They fail to capture informal group activities, such as the neighborhood soccer league that only uses un-reservable sports fields.

VivaParks events were correlated with lower event attendance compared to weeks following and preceding the week of the event. The sample size is too small for conclusions of significance.

Sports and picnic reservations were not correlated to park attendance. Holidays did not affect the relationship.

Amenities

Based on prior findings, one hypothesis was that amenities attracted people to parks. However, this analysis found that out of 46 amenities at 193 parks, sports fields have the strongest correlation to yearly visits. These include handball, soccer, volleyball, bocce ball, and multi-use fields as some of the most highly correlated. Handball and bocce ball courts cannot be reserved.

Additionally, other highly correlated amenities include games (such as picnic tables with a checkerboard on top), amphitheaters, and public decorations/monuments. These amenities are open and easy to access.

The results support the hypothesis that group activities are correlated with foot traffic. Many of these amenities encourage the gathering of small groups and cannot be reserved.

Additionally, there is a weak correlation between the number of park amenities and park visits. When looking at moderately popular parks, the number of amenities explains 10% of the variation of park visits. When looking at broadly popular, less popular, and locally popular parks, the number of amenities explains 42% of the variance in park visits. This indicates that there seems to be different patterns driving park use. Usage variance can only be partially explained by amenities.

Factors Combined

Finally, the City looked at the two main factors – area of park and total amenities at park – and their relationships to park visitations using a linear regression model. The number of events, picnics, and sports reservations were not included for two reasons. First, there was no linear relationship between these organized events and park visitations. Second, not many parks had events, picnics, and sports reservations, so there was too little sample size to cover all parks.

These factors explain much more of the variation in park visits for broadly, locally, and less popular parks than for moderately popular parks. The number of amenities at the park play a much larger role than the size of the park for broadly popular park visits. Number of amenities and the size of the park play somewhat equal amounts of influence for moderately popular parks. However, the size of the park plays a large role for all parks, locally popular parks, and less popular parks when explaining variation in park visits.

Percentage of Variation in Log(Park Visits) explained by factors	Log(Acres of Park) + Log(Total Amenities at Park)	Log(Total Amenities at Park)	Log(Acres of Park)
All Parks	18.8%	7.3%	14.3%
Broadly popular parks	11.3%	10.3%	0.0%
Locally popular parks	34.5%	8.0%	22.1%
Less popular parks	60.2%	7.0%	38.2%
Moderately popular parks	25.7%	13.5%	16.3%

Figure 5: Park size is influential in explaining park visitation variation for most parks except for broadly popular parks. Total amenities at parks is more influential when explaining moderately popular and broadly popular park visits.

A linear regression model created using the total amenities at the park to see the effect of total yearly visits to parks. This factor is most under PRNS's control.

For all parks, broadly popular and locally popular parks would benefit the most from adding amenities to the parks.

Park Type	Percent increase in park visits by 1% increase in total amenities
All Parks	2.7%
Broadly popular parks	5.8%
Locally popular parks	4.7%
Less popular parks	3.3%
Moderately popular parks	2%

Figure 6: Increasing amenities impact park visits at broadly popular and locally parks the most.

Conclusion:

The goal was to identify factors that drive people to parks and which factors are most worthy of investment. The main factors include the size of the park, total amenities at the park, number of events, number of picnics, and number of sports reservations. Demographic factors, such as race or age, are not associated with different kinds of parks. However, income does have some relationship with park visitors – park goers tend to make less income than the average San José resident.

As a result, this analysis recommends a focus on amenities to enable group activities, like sports and board games. In particular, sports amenities, game areas, and large gathering sites for groups of people to use as they wish are most highly correlated with park attendance.

Appendix:

1. **Broadly popular parks (red)**: parks that received over 100% of expected visitors based on the number of CBGs that visited the park and had more than 400 CBGs visit the park in a year.
    a. The 22 parks are: Alum Rock Park, Backesto Park, Brigadoon Tot Lot, Cataldi Park, Emma Prusch Farm Park, Flickinger Park, George Page Park, Gran Paradiso Park, Guadalupe Gardens, Kelley Park, Lake Cunningham Park, Moitozo Park, Murdock Park, O'Connor Park, Parkview II Park, Plaza de Cesar Chavez, Ramac Park, Starbird Park, St. James Park, Watson Park, William Street Park, Wilcox Park
2. **Locally popular parks (blue)**: parks that received at least 90% of expected visitors based on the number of CBGs that visited the park and had less than 150 CBGs visit the park in a year.
    a. The 24 parks are: Basking Ridge Park, Bestor Art Park, Bonita Park, Chris Hotts Park, Elaine Richardson Park, Fleming Park, Forestdale Tot Lot, Iris Chang Park, Lone Bluff Park, McLaughlin Park, Meadows Park, Melody Park, Metcalf Park, O'Donnell's Gardens Park, Parkview I Park, Parkview III Park, Parque de la Amistad,

Parque de los Pobladores, Playa del Rey Park, Pfeiffer Park, Shady Oaks Park, Turtle Rock Park

3. **Less popular parks (green)**: parks that received less than 10% of expected visitors based on the number of CBGs that visited the park and had less than 50 CBGs visit the park in a year.
    a. The 26 parks are: Bellevue Park, Brenda Lopez Memorial Plaza, Cahill Park, Canyon Creek Park, Carolyn Norris Park, Century Oaks Park, Dove Hill Park, Foothill Park, Fuller Avenue Park, Gleason Park, Gregory Tot Lot, Hester Park, Meadowfair Park, Nancy Lane Plaza, Nisich Park, Roberto Antonio Balermino Park, Rubino Park, Ryland Dog Park, San Antonio Tot Lot, Theodore Lenzen Park, Thousand Oaks Park, Vinci Park, William H. Cilker Park, William Lewis Manly Park, Windmill Springs Park, Zolezzi Park
4. **Moderately popular parks (yellow):** these parks performed as expected and did not fall into any of the prior groups.
    a. The 131 parks in this group are: Aborn Park, Avenida Espana Park, Arcadia Ballpark, Baypointe Interim Park, Berryessa Creek Park, Boggini Park, Branham Park, Brigadoon Park, Brooktree Park, Butcher Park, Buena Vista Park, Cahalan Park, Calabazas Park, Calero Park, Canoas Park, Cannery Park, Capitol Park, Carrabelle Park, Cathedral Oaks Park, Charlotte Commons, Chynoweth Park, Cimarron Park, Commodore Park, Comanche Park, Danna Rock Park, De Anza Park, Del Monte Park, Discovery Dog Park, Discovery Meadow, Doerr Park, Erikson Park, Esther Medina Park, Falls Creek Park, Fernish Park, Fowler Creek Park, Frank M. Santana Park, Glenview Park, Great Oaks Park, Greystone Park, Guadalupe Gardens - Courtyard Garden, Guadalupe Gardens - Heritage Rose Garden, Guadalupe Gardens - Historic Orchard, Guadalupe Oak Grove Park, Guadalupe River Park - Arena Green, Guadalupe River Park - Arena Green (East Confluence), Guadalupe River Park - Discovery Meadow, Gullo Park, Groesbeck Hill Park, Hathaway Park, Hacienda Park, Houge Park, Hummingbird Park, Jeffery Fontana Park, John Mise Park, John P. McEnery Park, Kelley Park - Happy Hollow Park and Zoo, Kelley Park - Japanese Friendship Garden, Kelley Park - Vietnamese Heritage Garden, Kirk Park, La Colina Park, La Ragione Tot Lot, Lone Hill Park, LoBue Park, Luna Park, Madden Park, Marijane Hamann Park, Martin Park, Mayfair Park, Mineta Park, Miner Park, Miyuki Dog Park, Monta Vista Park, Montgomery Hill Park, Mt. Pleasant Park, Municipal Rose Garden, Newhall Park, Northwood Park, Norwood Creek Park, Orchard Park, Overfelt Gardens, Palmia Park, Parma Park, Parque de Padre Mateo Sheedy, Paul Moore Park, Piercy Park, Plata Arroyo Park, Rainbow Park, Raleigh Park, Ramblewood Park, Raymond Bernal Jr. Memorial Park, Richard E Huerta Park, Rincon South Park, River Glen Park, River Oaks Park, Riverbend Park, Rocksprings Park, Roosevelt Park, Rosemary Gardens Park, Roy Avenue Park, Ryland Park, Russo Park, San Tomas Park, Saratoga Creek Park, Scottsdale Park, Selma Olinder Park, Silver Creek Linear Park, Silver Leaf Park, Solari Park, Stonegate Park, St.

Elizabeth Park, Sylvia Cassell Park, Tamien Park, T.J. Martin Park, Terrell Park, Townsend Park, Vieira Park, Vieira Park Overlook, Vista Montaña (1 acre), Vista Montaña (5 acre), Vista Park, Wallenberg Park, Waterford Park, West Evergreen Park, Willow Street Frank Bramhall Park