

datahack_geomap

Agneya Poduval

2025-05-27

Load Dataset

```
options(tigris_use_cache = TRUE)

# Load rent data
rent_data <- read_csv("clean_california_zillow_rent.csv") %>%
  filter(!is.na(RentIndex)) %>%
  mutate(RegionName = str_pad(as.character(ZIPCode), 5, pad = "0"))

# Download and simplify ZIP shapes (for CA ZIPs)
zips_needed <- unique(rent_data$RegionName)
zip_shapes <- zctas(cb = FALSE, starts_with = "9") |>
  filter(ZCTA5CE20 %in% zips_needed) |>
  st_transform(3310) |>
  st_simplify(dTolerance = 100) # Simplify to reduce memory

# Join rent data
rent_sf <- left_join(zip_shapes, rent_data, by = c("ZCTA5CE20" = "RegionName"))

# UC campuses
uc_campuses <- tribble(
  ~Campus, ~Longitude, ~Latitude,
  "UC Berkeley", -122.2585, 37.8719,
  "UC Davis", -121.7446, 38.5382,
  "UC Irvine", -117.8443, 33.6405,
  "UCLA", -118.4452, 34.0689,
  "UC Merced", -120.4189, 37.3649,
  "UC Riverside", -117.3281, 33.9737,
  "UC San Diego", -117.2394, 32.8801,
  "UC San Francisco", -122.4580, 37.7631,
  "UC Santa Barbara", -119.8489, 34.4139,
  "UC Santa Cruz", -122.0622, 36.9916
)

uc_sf <- st_as_sf(uc_campuses, coords = c("Longitude", "Latitude"), crs = 4326) |>
  st_transform(3310)
```

Extra

```
# Extract year
rent_data <- rent_data %>%
  filter(!is.na(RentIndex)) %>%
  mutate(RegionName = str_pad(as.character(ZIPCode), 5, pad = "0"),
         Year = lubridate::year(as.Date(Date)))

# Separate datasets
rent_2015 <- rent_data %>% filter(Year == 2015)
rent_2025 <- rent_data %>% filter(Year == 2025)

# Join ZIPs with rent data
rent_sf_2015 <- left_join(zip_shapes, rent_2015, by = c("ZCTA5CE20" = "RegionName"))
rent_sf_2025 <- left_join(zip_shapes, rent_2025, by = c("ZCTA5CE20" = "RegionName"))

for (i in 1:nrow(uc_sf)) {
  campus_name <- uc_sf$Campus[i]
  campus_pt <- uc_sf[i, ]

  # Buffers
  buffer_1 <- st_buffer(campus_pt, 1609)
  buffer_3 <- st_buffer(campus_pt, 1609 * 3)
  buffer_5 <- st_buffer(campus_pt, 1609 * 5)

  # Nearby ZIPs
  zip_2015 <- rent_sf_2015[st_intersects(rent_sf_2015, buffer_5, sparse = FALSE), ]
  zip_2025 <- rent_sf_2025[st_intersects(rent_sf_2025, buffer_5, sparse = FALSE), ]

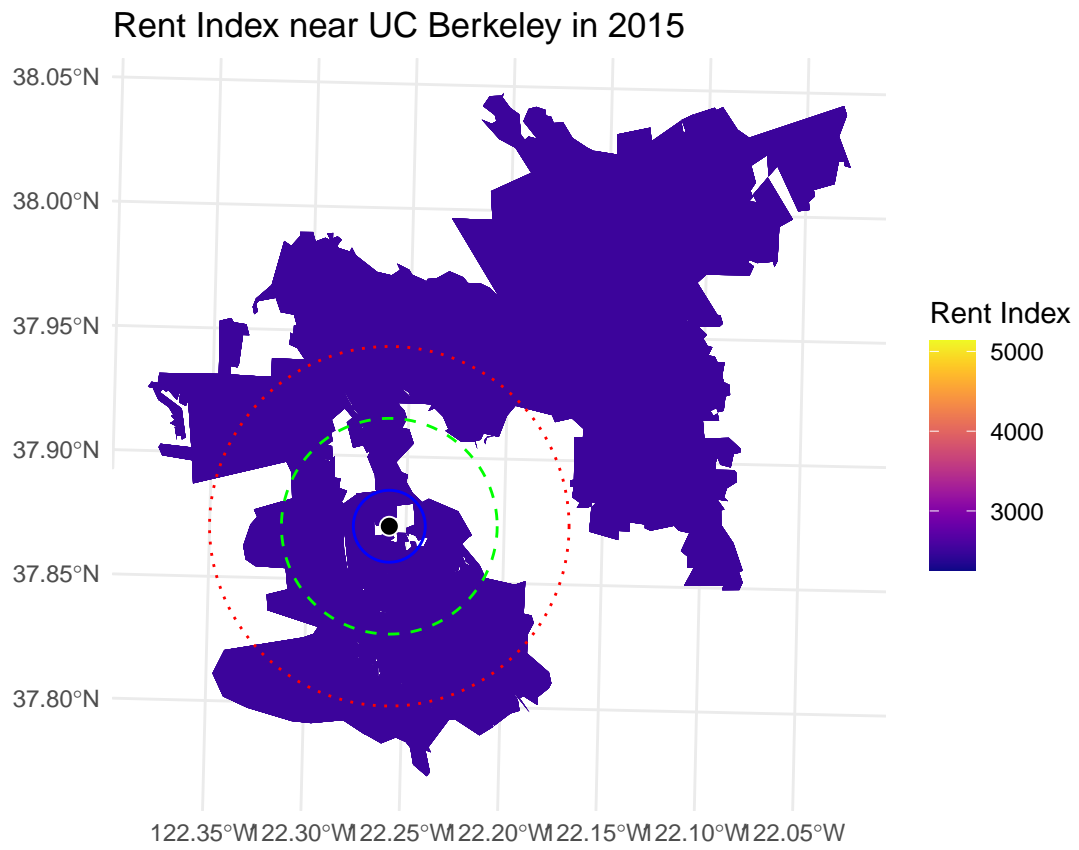
  # Campus-specific scale
  campus_min <- min(c(zip_2015$RentIndex, zip_2025$RentIndex), na.rm = TRUE)
  campus_max <- max(c(zip_2015$RentIndex, zip_2025$RentIndex), na.rm = TRUE)

  for (year in c(2015, 2025)) {
    zip_year <- if (year == 2015) zip_2015 else zip_2025

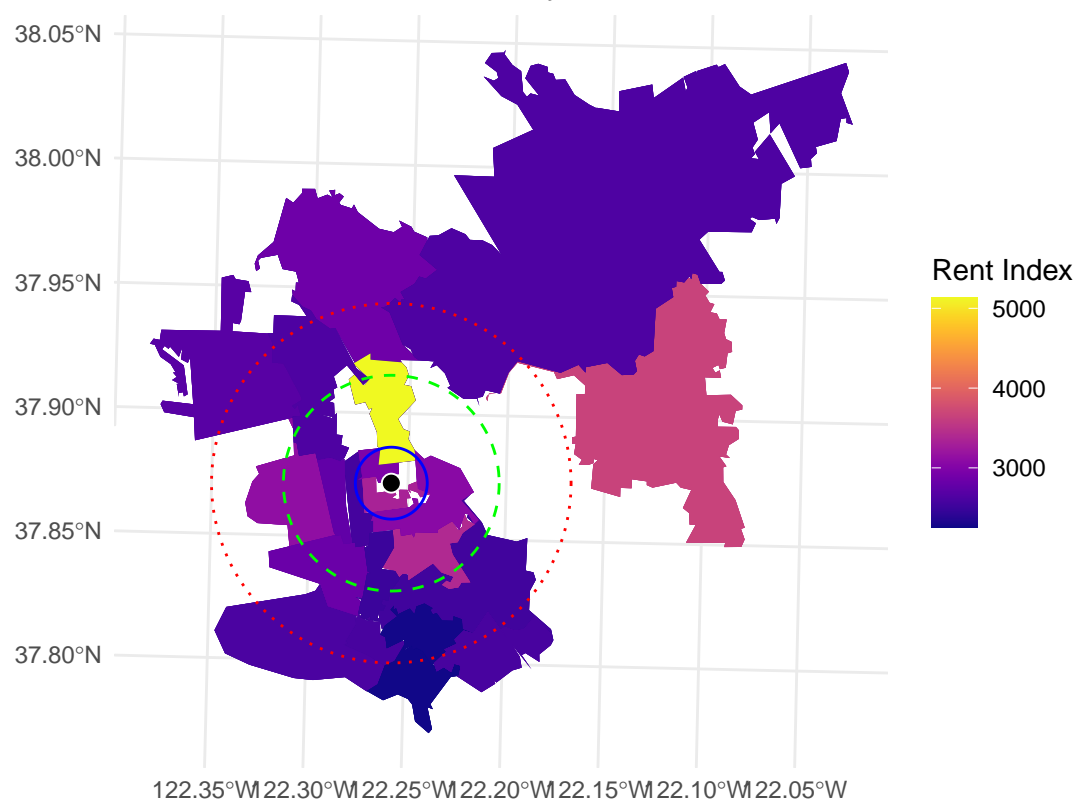
    p <- ggplot() +
      geom_sf(data = zip_year, aes(fill = RentIndex), color = NA) +
      scale_fill_viridis(
        option = "plasma",
        name = "Rent Index",
        limits = c(campus_min, campus_max),
        na.value = "grey90"
      ) +
      geom_sf(data = buffer_5, fill = NA, color = "red", linetype = "dotted", linewidth = 0.5) +
      geom_sf(data = buffer_3, fill = NA, color = "green", linetype = "dashed", linewidth = 0.5) +
      geom_sf(data = buffer_1, fill = NA, color = "blue", linetype = "solid", linewidth = 0.5) +
      geom_sf(data = campus_pt, shape = 21, fill = "black", color = "white", size = 3) +
      labs(title = paste0("Rent Index near ", campus_name, " in ", year)) +
      theme_minimal()

    print(p)
  }
}
```

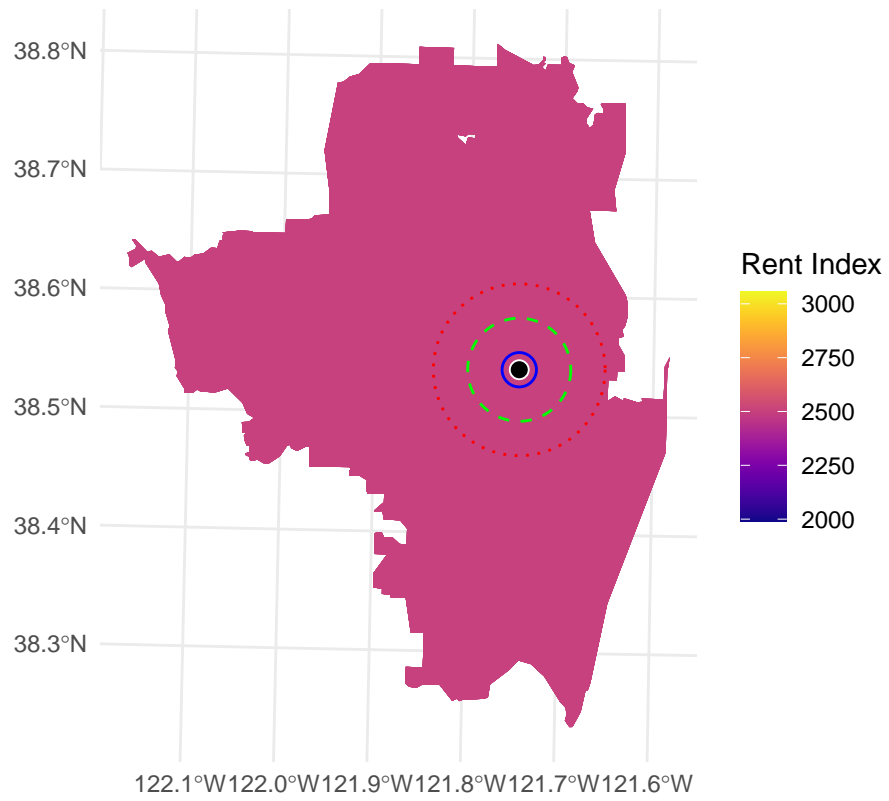
}



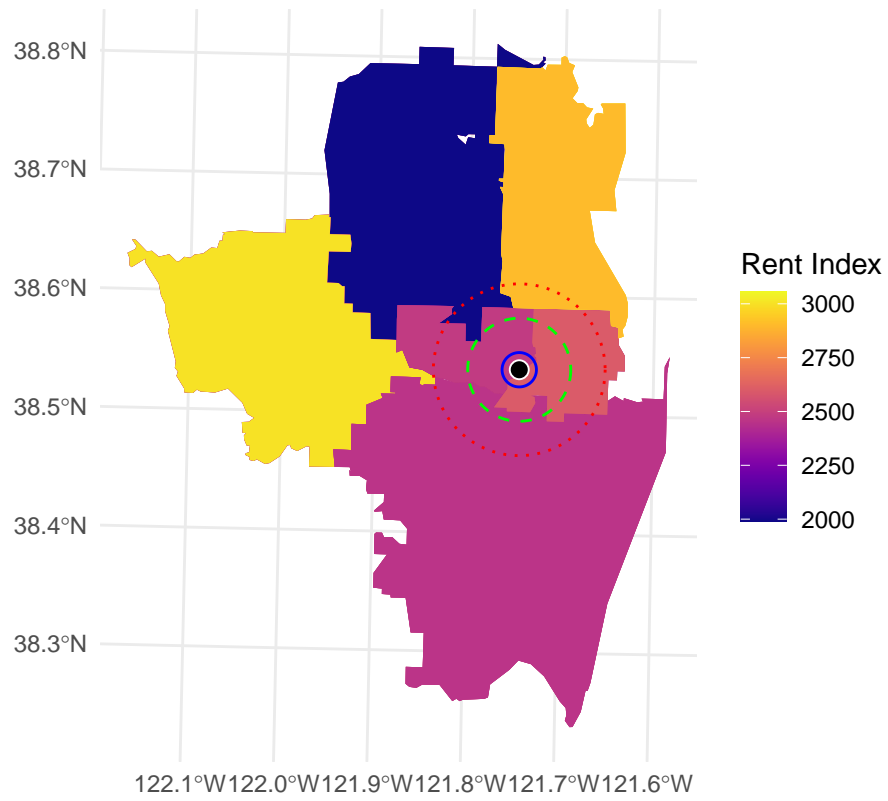
Rent Index near UC Berkeley in 2025



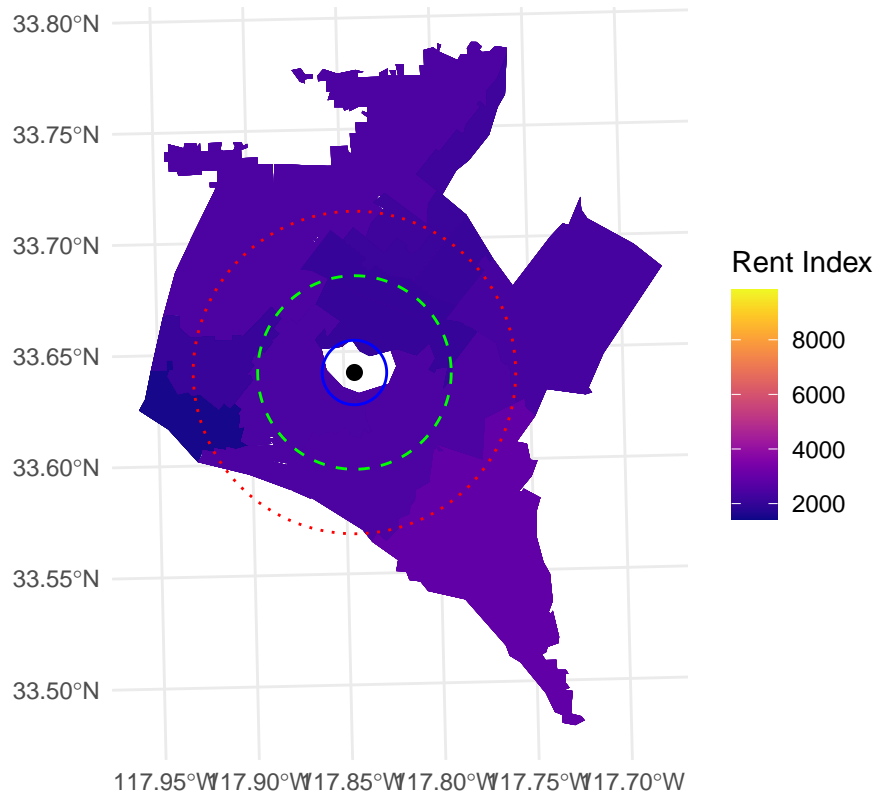
Rent Index near UC Davis in 2015



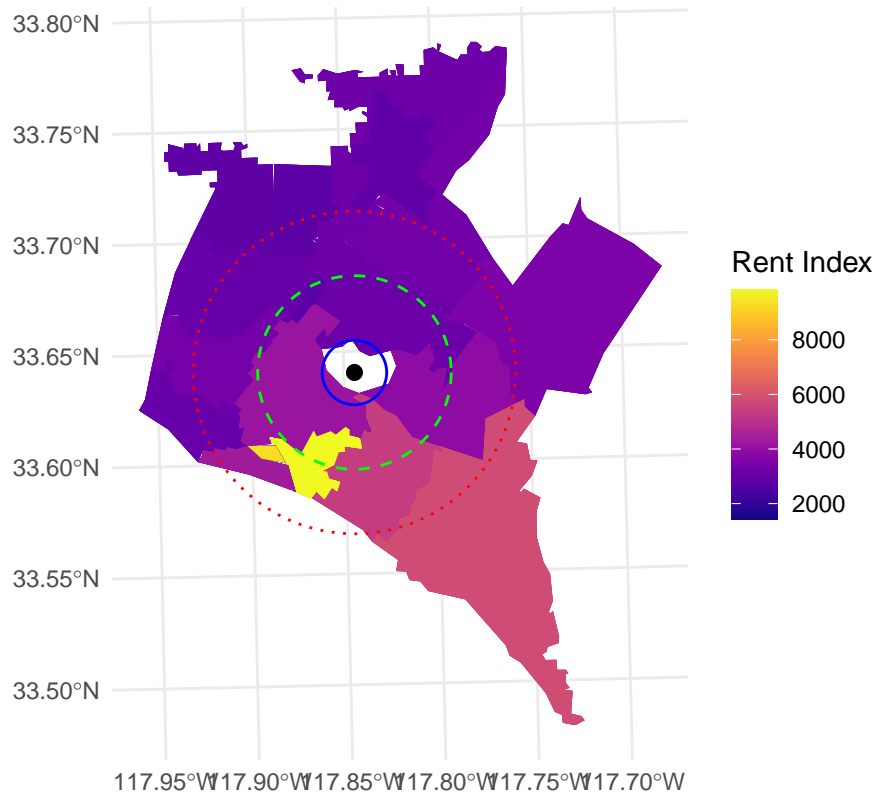
Rent Index near UC Davis in 2025



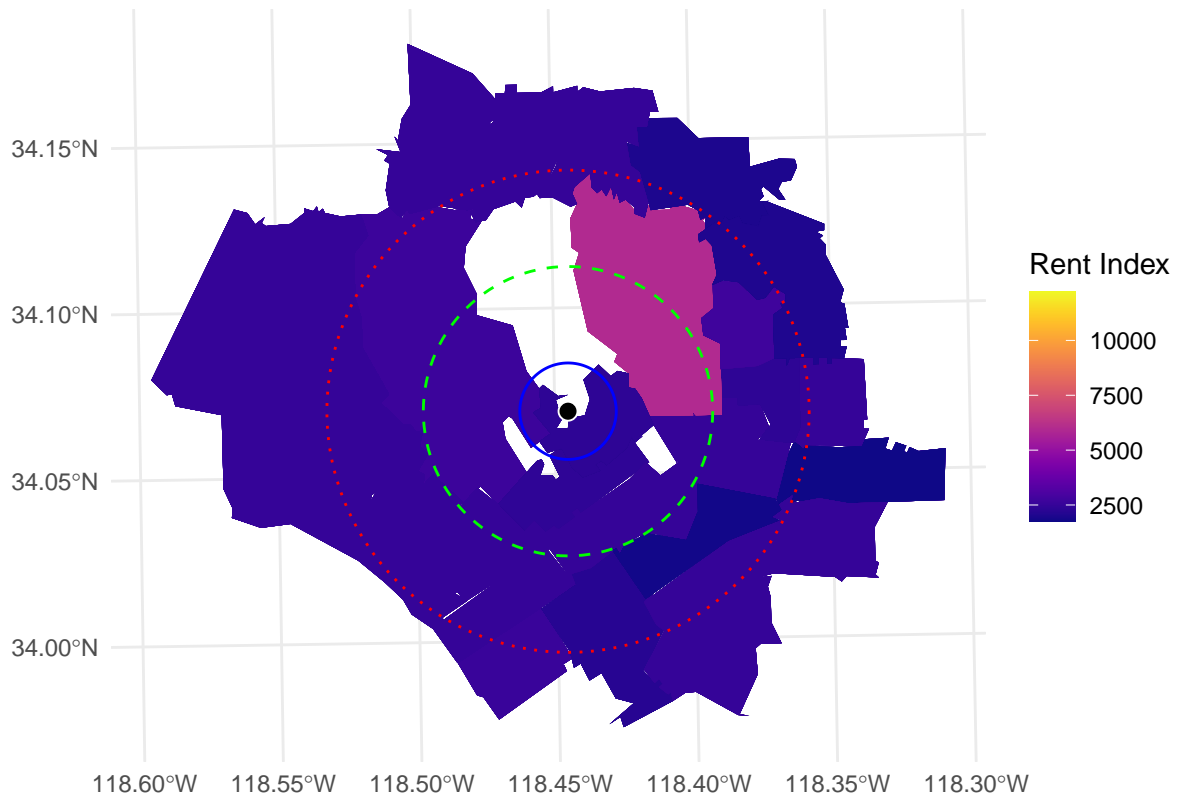
Rent Index near UC Irvine in 2015



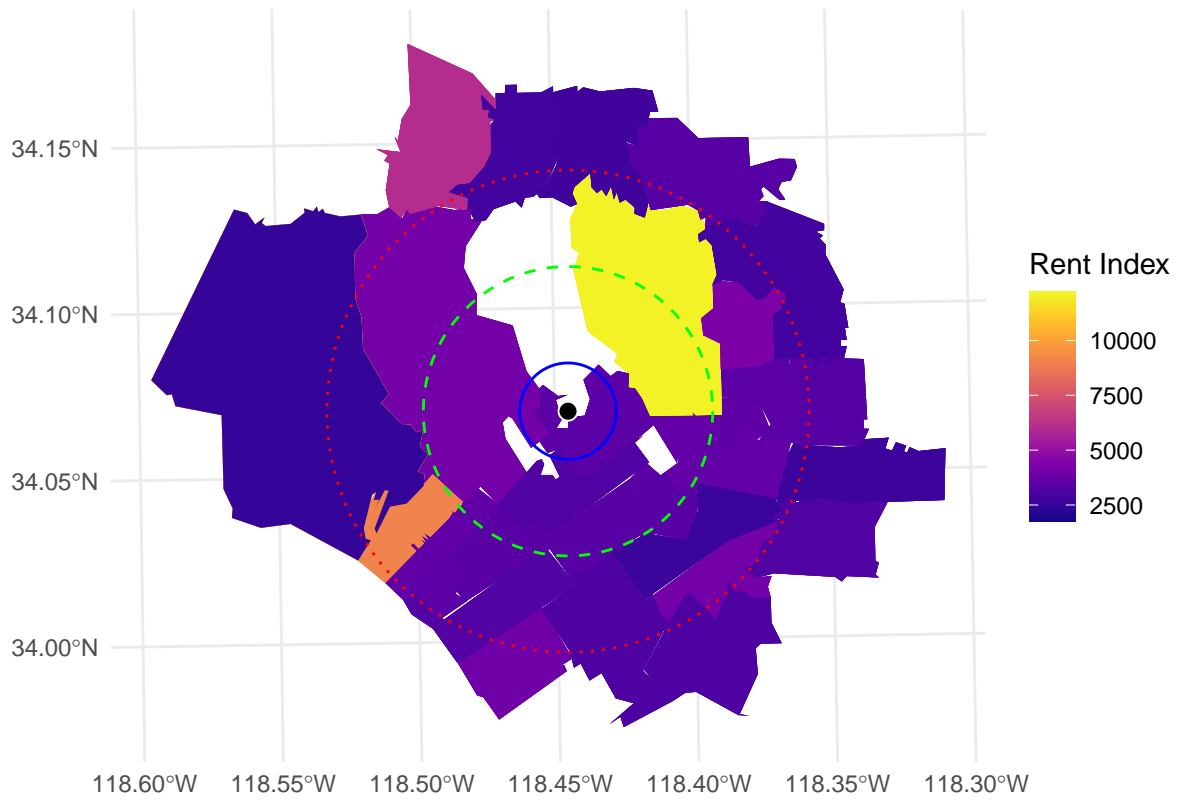
Rent Index near UC Irvine in 2025



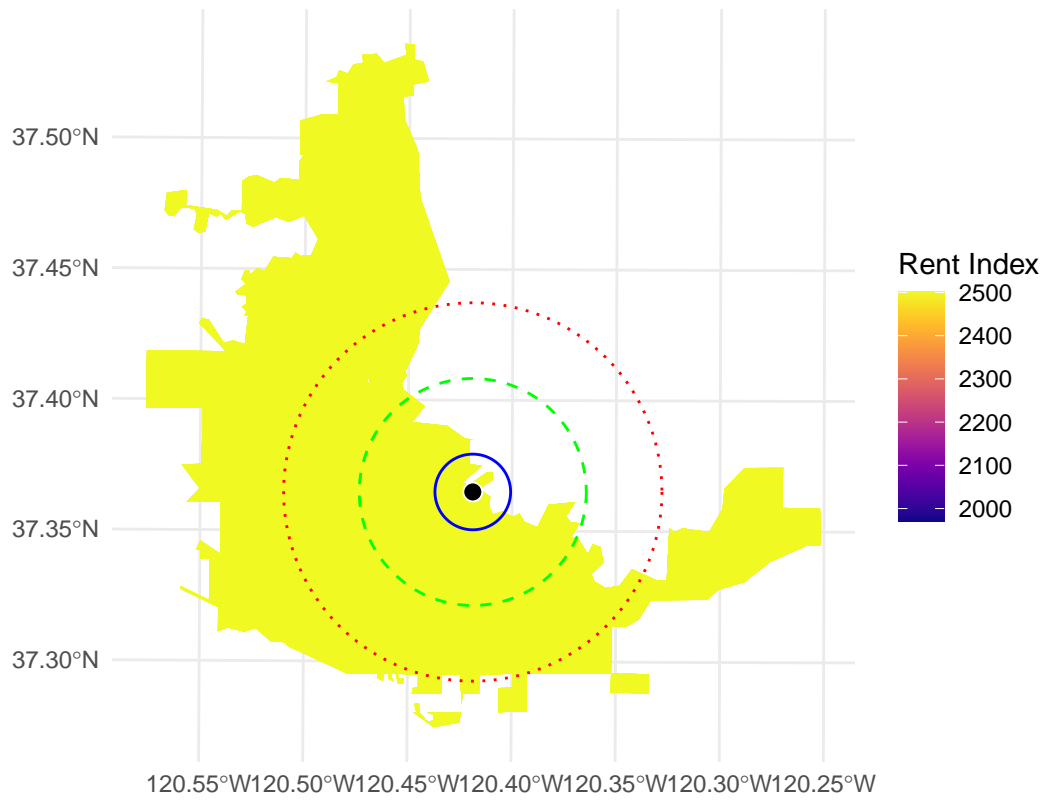
Rent Index near UCLA in 2015



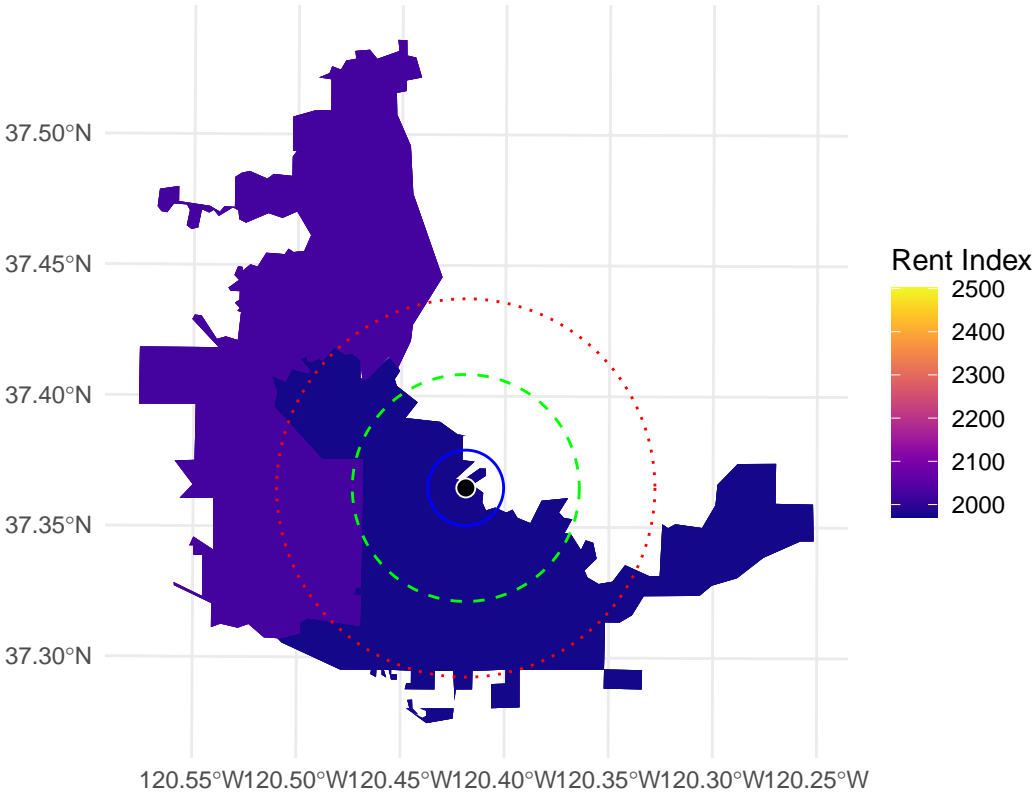
Rent Index near UCLA in 2025



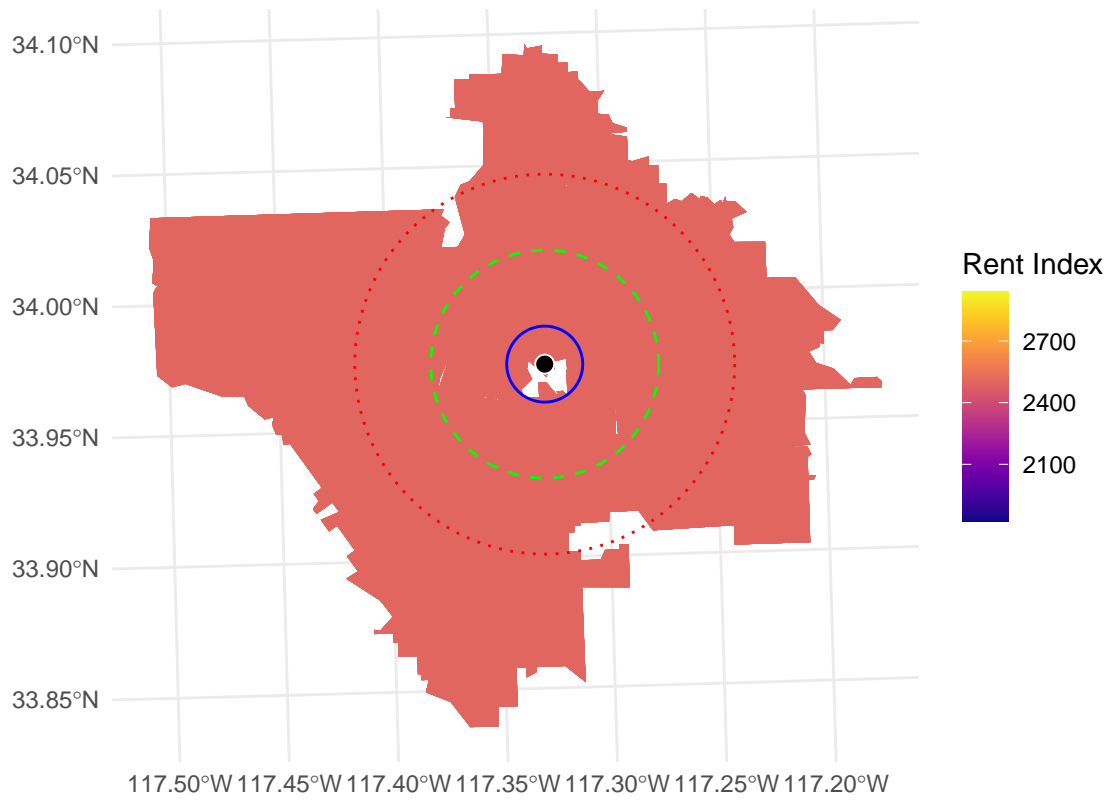
Rent Index near UC Merced in 2015



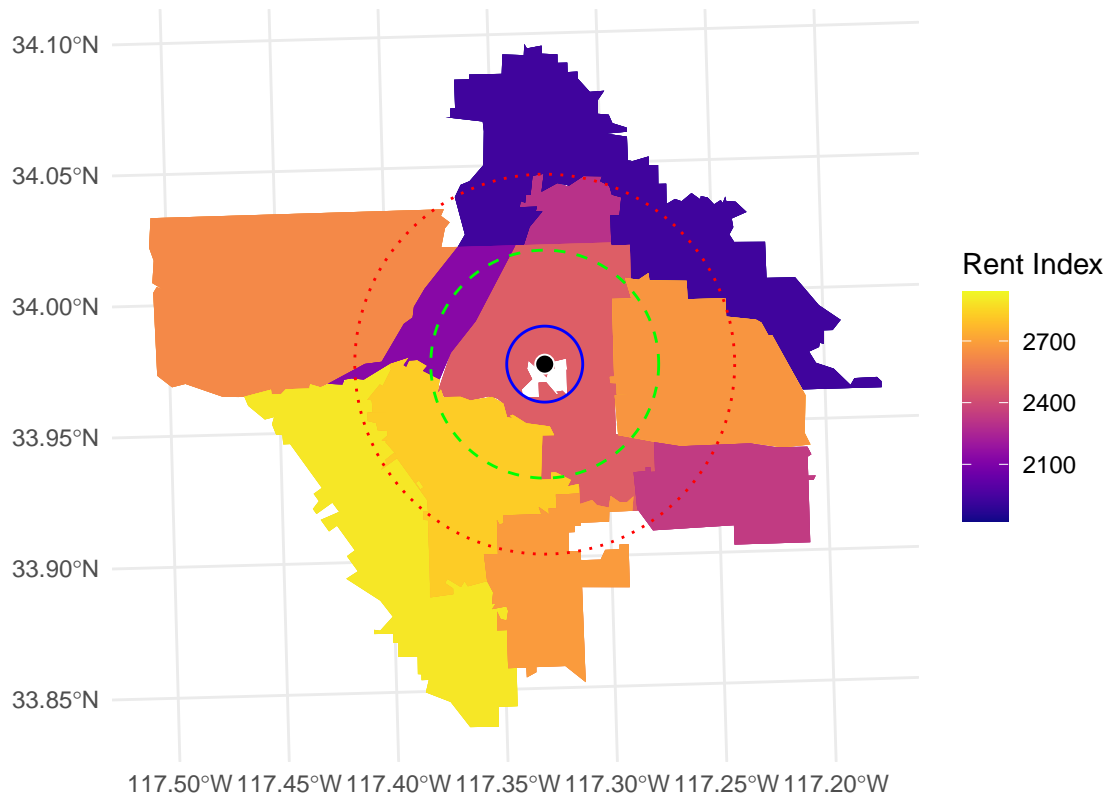
Rent Index near UC Merced in 2025



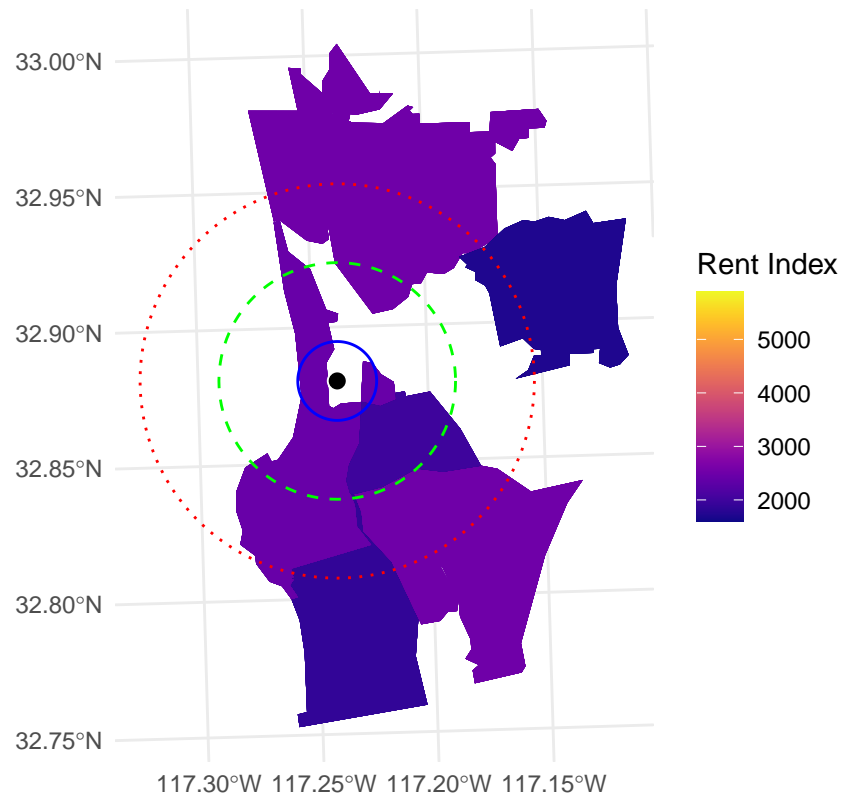
Rent Index near UC Riverside in 2015



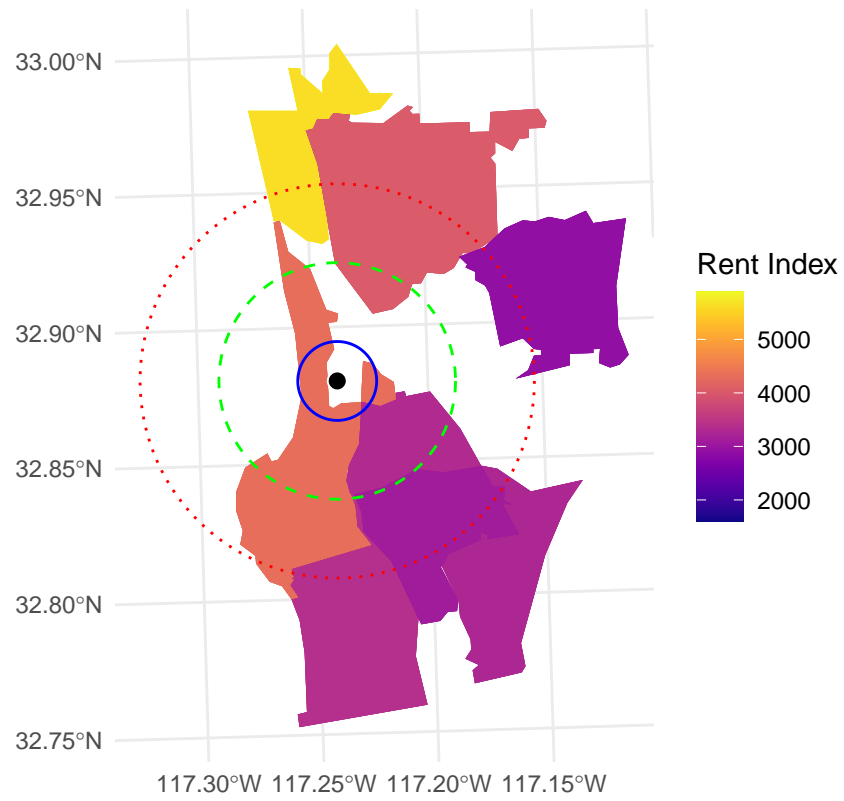
Rent Index near UC Riverside in 2025



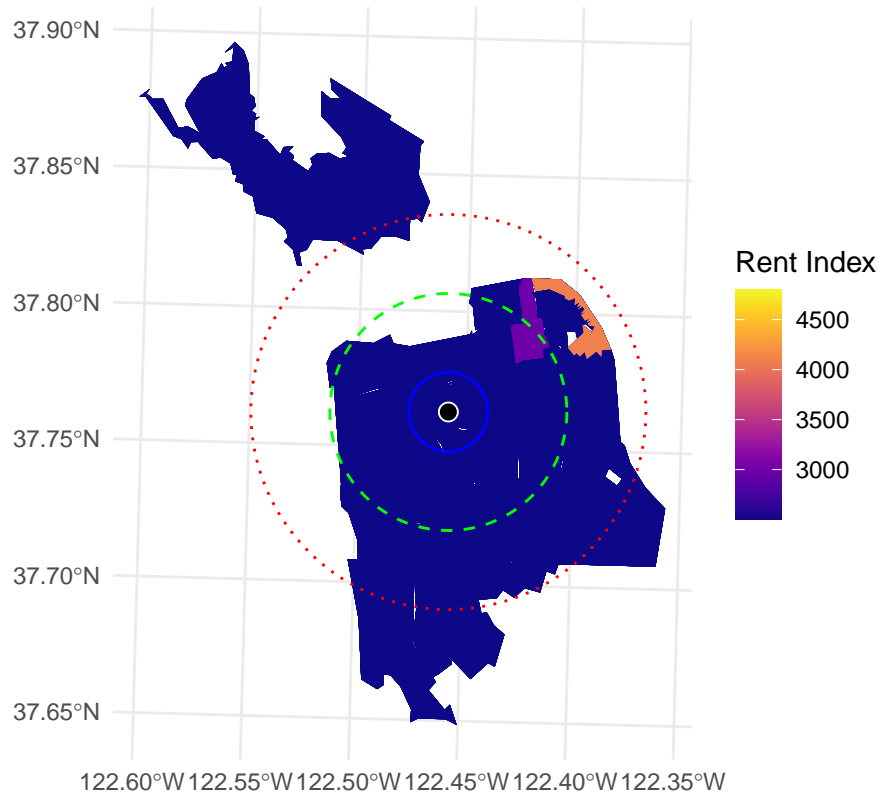
Rent Index near UC San Diego in 2015



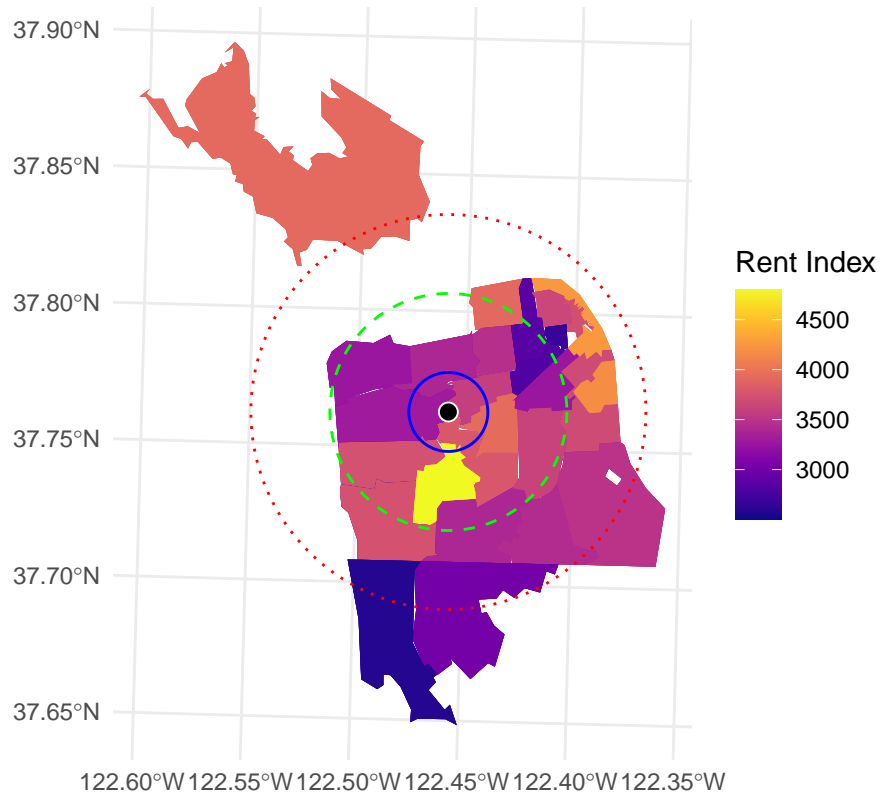
Rent Index near UC San Diego in 2025



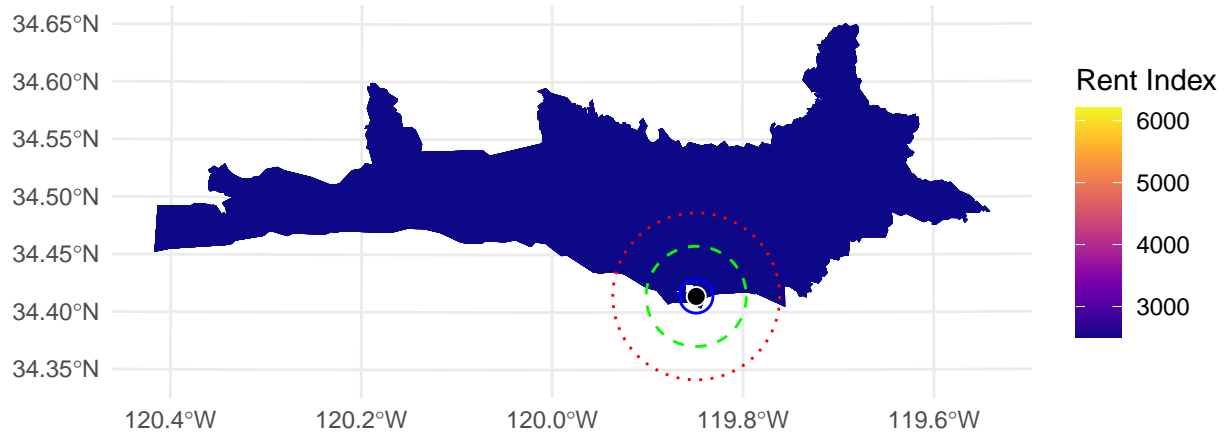
Rent Index near UC San Francisco in 2015



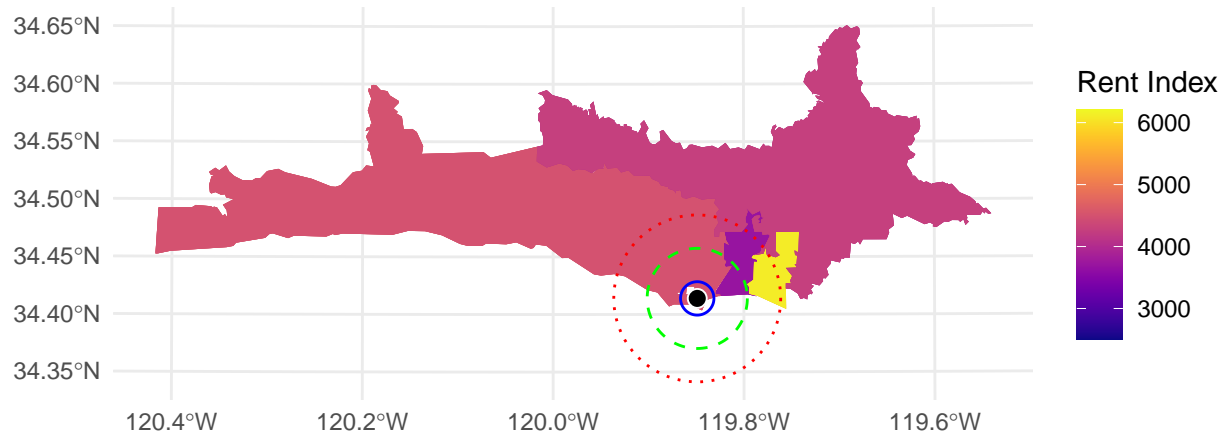
Rent Index near UC San Francisco in 2025



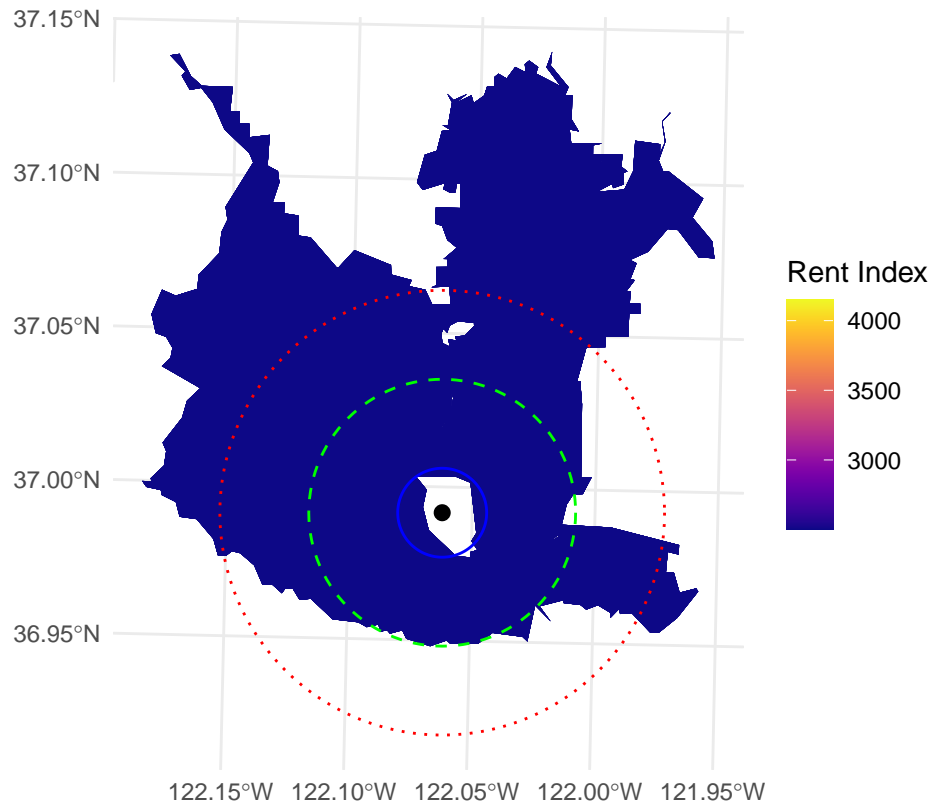
Rent Index near UC Santa Barbara in 2015

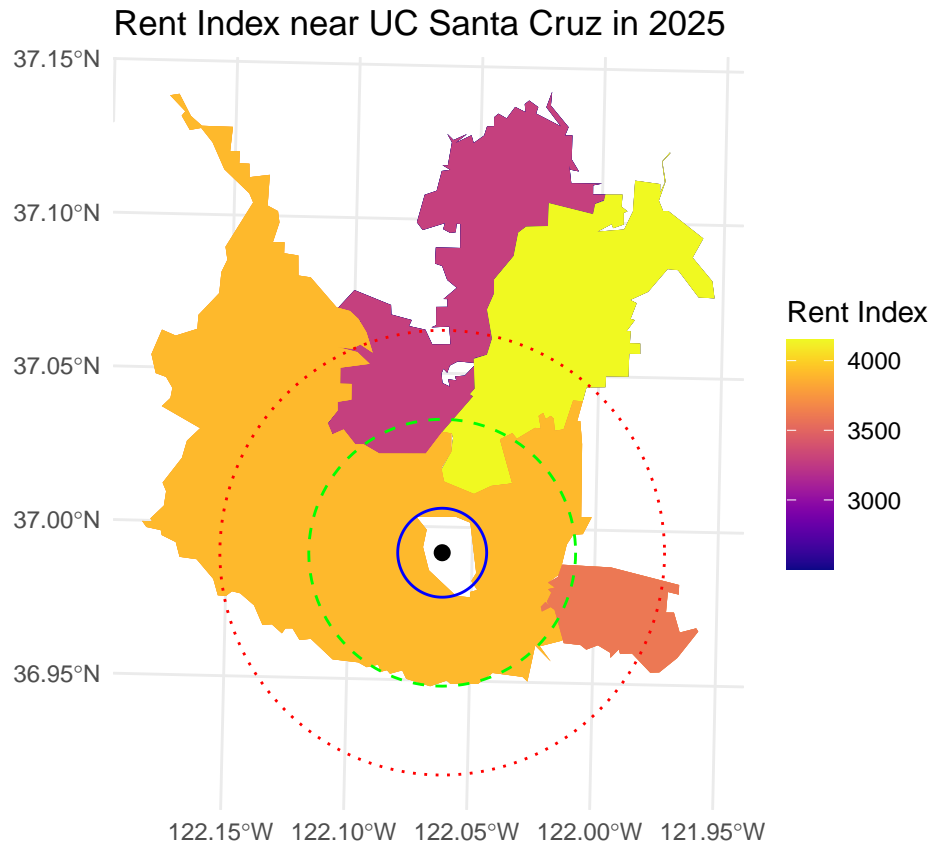


Rent Index near UC Santa Barbara in 2025



Rent Index near UC Santa Cruz in 2015





Adjusted for inflation and rent increase

```
rent_data <- read_csv("clean_california_zillow_rent.csv") %>%
  filter(!is.na(RentIndex)) %>%
  mutate(ZIPCode = str_pad(as.character(ZIPCode), 5, pad = "0"))

# Step 1: Separate 2015 and 2025 data
rent_2015 <- rent_data %>%
  filter(Date == "2015-01-31") %>%
  select(ZIPCode, RentIndex_2015 = RentIndex)

rent_2025 <- rent_data %>%
  filter(Date == "2025-01-31") %>%
  select(ZIPCode, RentIndex_2025 = RentIndex)

# Step 2: Join the datasets by ZIP code
rent_joined <- inner_join(rent_2015, rent_2025, by = "ZIPCode") %>%
  mutate(
    RentIndex_2025_adj = RentIndex_2025 / (1.05 ^ 10),
    Expected_2025 = RentIndex_2015 * (1.05 ^ 10),
    Diff_vs_expected = RentIndex_2025 - Expected_2025
  )
```

```

)
zip_shapes <- zctas(cb = FALSE, starts_with = "9") %>%
  st_transform(3310)

# Check ZIP format
zip_shapes$ZCTA5CE20 <- as.character(zip_shapes$ZCTA5CE20)

common_zips <- intersect(unique(zip_shapes$ZCTA5CE20), unique(rent_joined$ZIPCode))
length(common_zips)

## [1] 898

head(common_zips)

## [1] "94587" "95126" "90602" "95130" "95035" "92657"

# Step 3: Adjust 2025 rent for inflation (CPI from BLS or estimate)
cpi_2015 <- 233.707
cpi_2025 <- 317.67
inflation_adjustment <- 1 + ((cpi_2025 - cpi_2015) / cpi_2015)
rent_joined <- rent_joined %>%
  mutate(RentIndex_2025_adj = RentIndex_2025 * inflation_adjustment)

# Step 4: Compute expected rent under 5% annual increase
years <- 2025 - 2015
growth_rate <- 0.05

rent_joined <- rent_joined %>%
  mutate(
    Expected_2025 = RentIndex_2015 * (1 + 0.05)^10 * inflation_adjustment,
    Diff_vs_expected = RentIndex_2025 - Expected_2025
  )

# Step 5: Join with spatial ZIPs
zip_shapes$ZCTA5CE20 <- as.character(zip_shapes$ZCTA5CE20)
rent_joined$ZIPCode <- as.character(rent_joined$ZIPCode)

# Join with ZIP shapefile
rent_diff_sf <- left_join(zip_shapes, rent_joined, by = c("ZCTA5CE20" = "ZIPCode"))

for (i in 1:nrow(uc_sf)) {
  campus_name <- uc_sf$Campus[i]
  campus_pt <- uc_sf[i, ]

  buffer_1 <- st_buffer(campus_pt, 1609)
  buffer_3 <- st_buffer(campus_pt, 1609 * 3)
  buffer_5 <- st_buffer(campus_pt, 1609 * 5)

  # Filter nearby ZIPs for joined inflation-adjusted data
  zip_nearby <- rent_diff_sf[st_intersects(rent_diff_sf, buffer_5, sparse = FALSE), ]

  # Plot actual vs expected difference

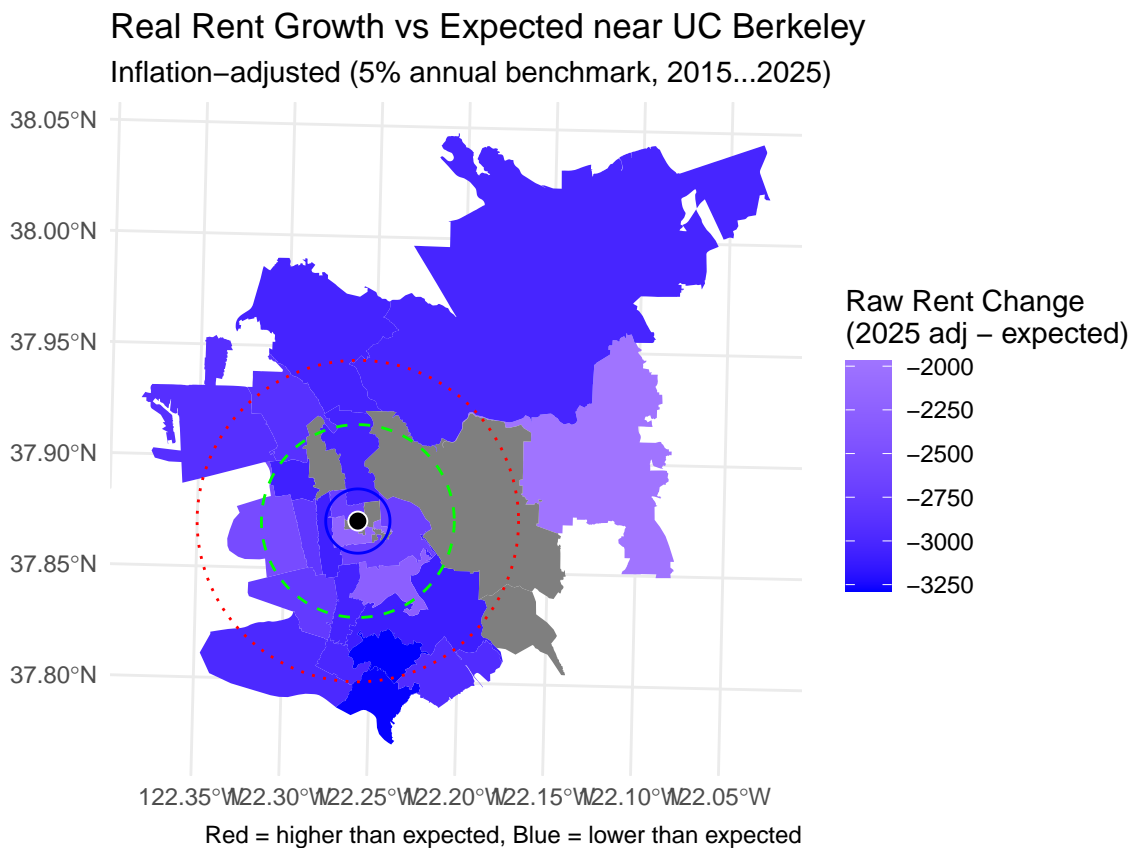
```

```

p <- ggplot() +
  geom_sf(data = zip_nearby, aes(fill = Diff_vs_expected), color = NA) +
  scale_fill_gradient2(
    low = "blue", mid = "white", high = "darkred",
    midpoint = 0, name = "Raw Rent Change\n(2025 adj - expected)"
  ) +
  geom_sf(data = buffer_5, fill = NA, color = "red", linetype = "dotted", linewidth = 0.5) +
  geom_sf(data = buffer_3, fill = NA, color = "green", linetype = "dashed", linewidth = 0.5) +
  geom_sf(data = buffer_1, fill = NA, color = "blue", linetype = "solid", linewidth = 0.5) +
  geom_sf(data = campus_pt, shape = 21, fill = "black", color = "white", size = 3) +
  labs(
    title = paste0("Real Rent Growth vs Expected near ", campus_name),
    subtitle = "Inflation-adjusted (5% annual benchmark, 2015-2025)",
    caption = "Red = higher than expected, Blue = lower than expected"
  ) +
  theme_minimal()

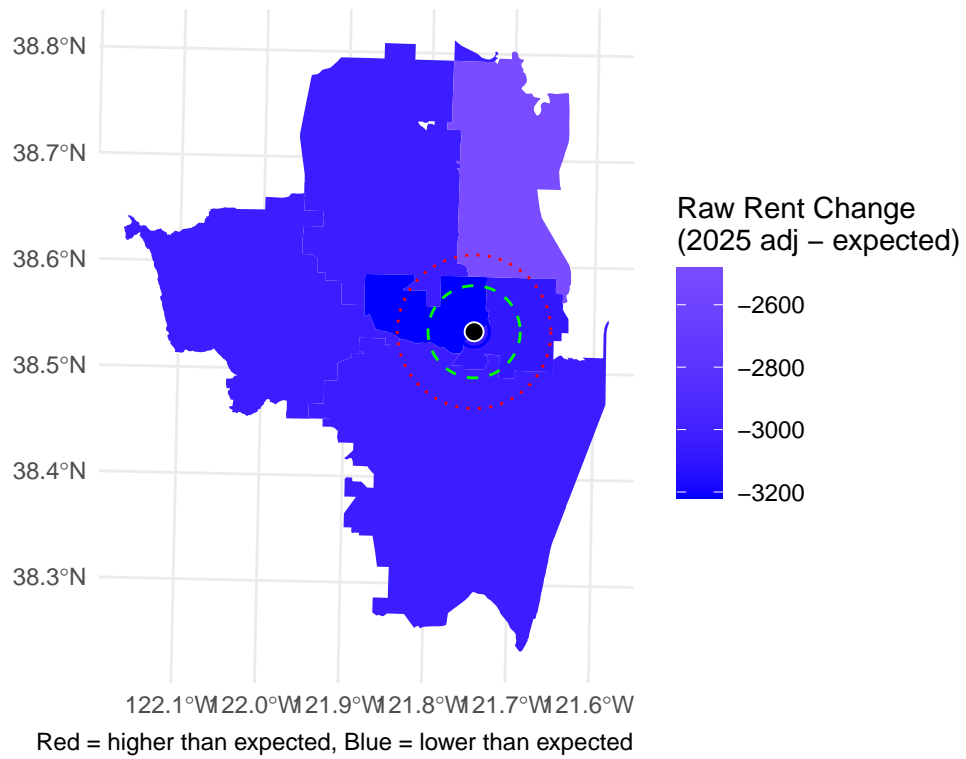
print(p)
}

```



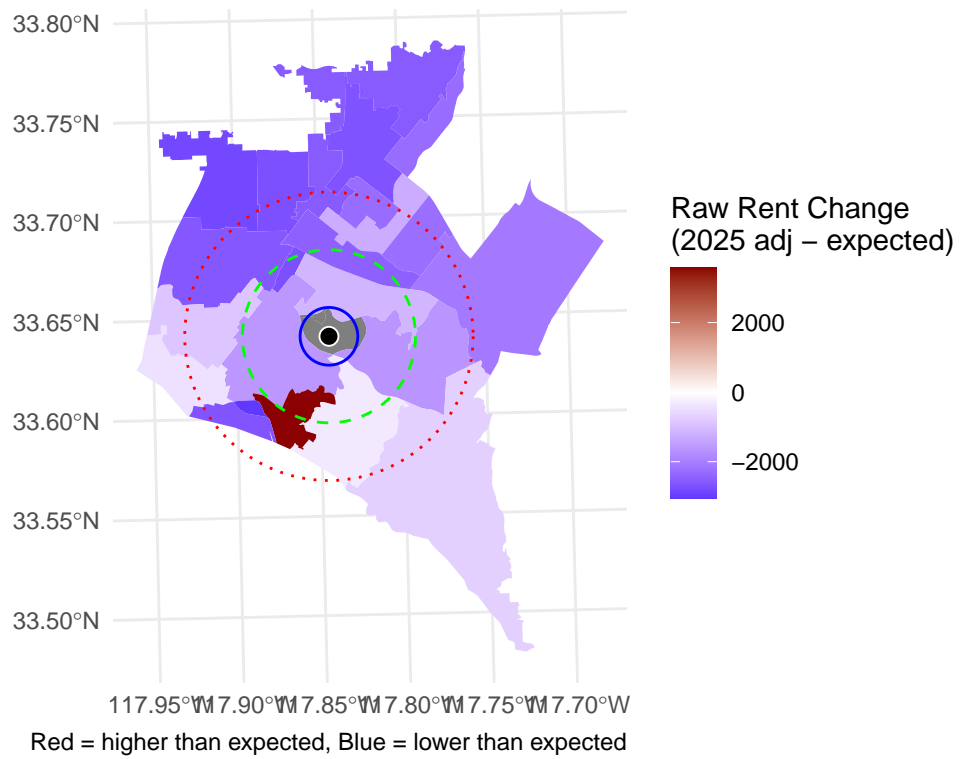
Real Rent Growth vs Expected near UC Davis

Inflation-adjusted (5% annual benchmark, 2015...2025)

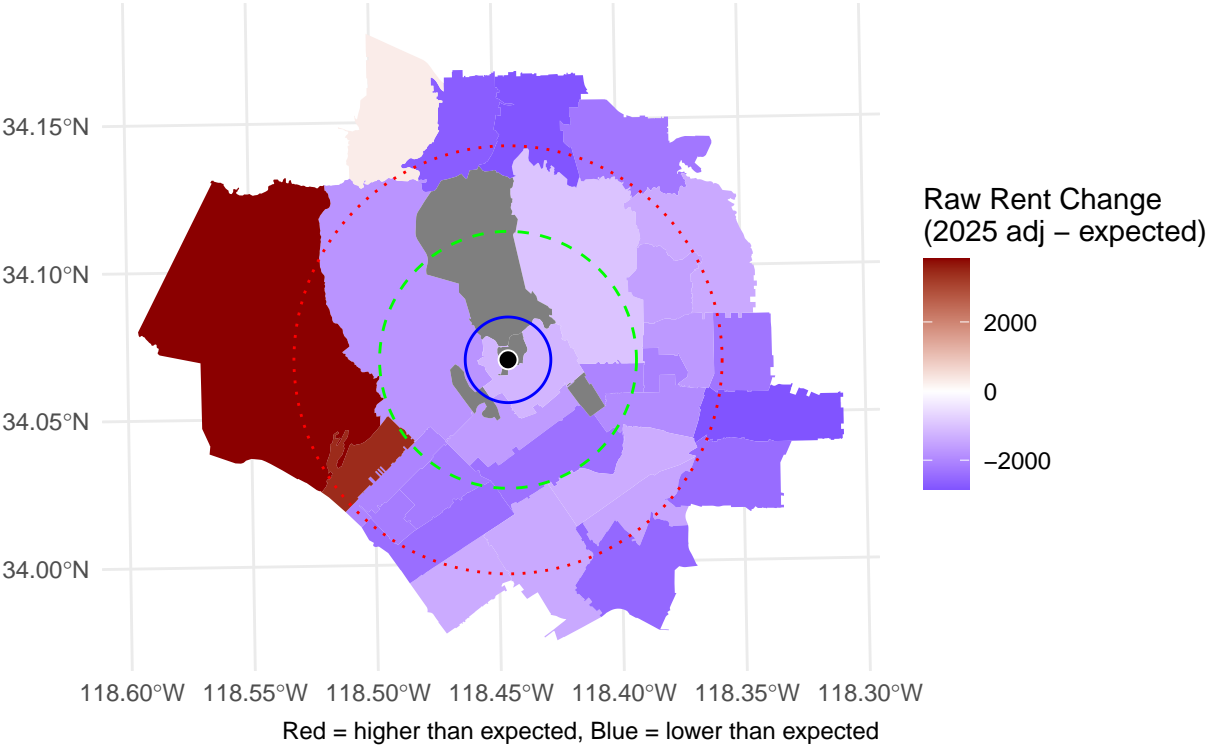


Real Rent Growth vs Expected near UC Irvine

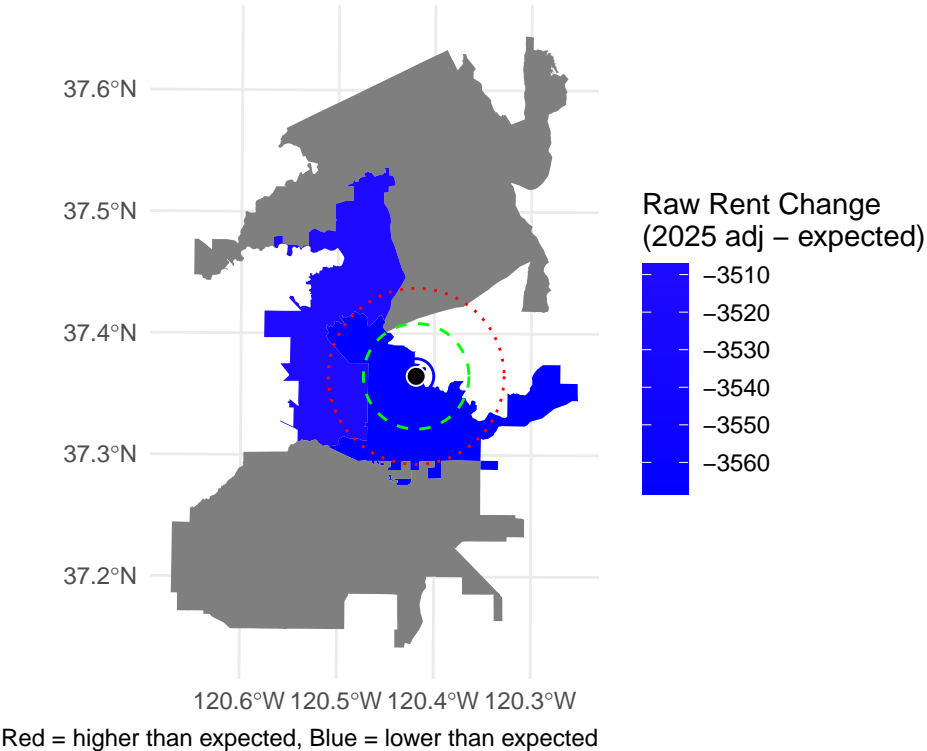
Inflation-adjusted (5% annual benchmark, 2015...2025)



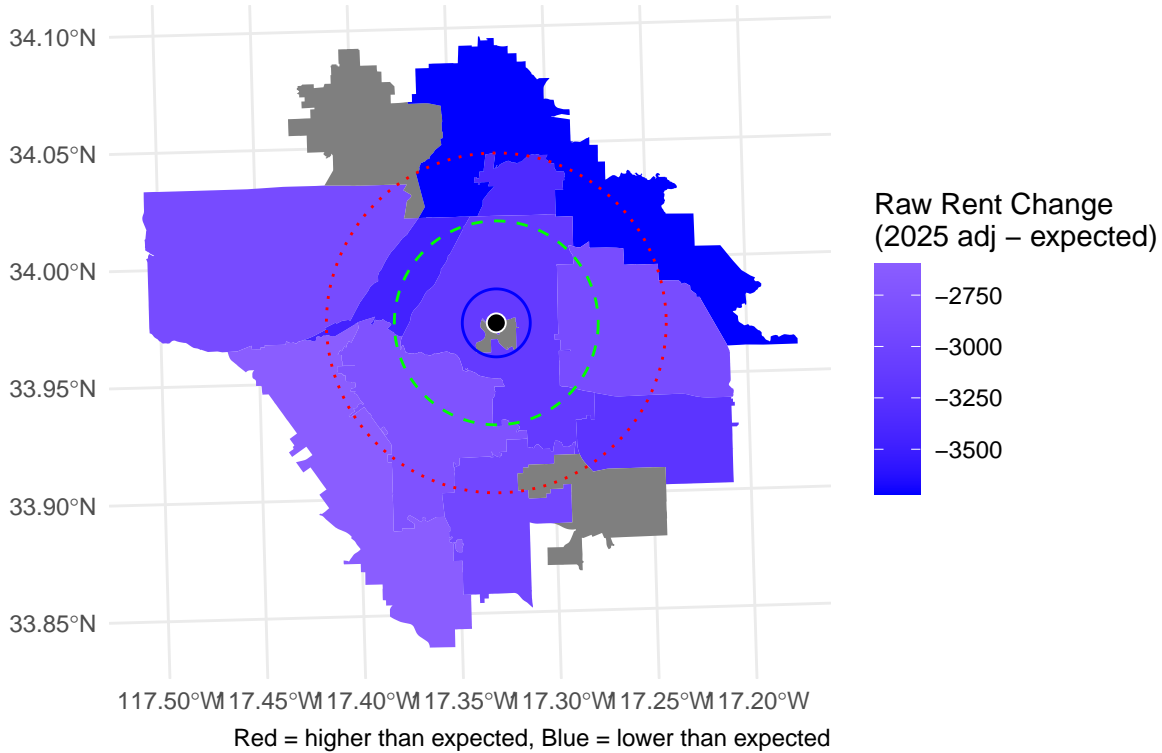
Real Rent Growth vs Expected near UCLA
Inflation-adjusted (5% annual benchmark, 2015...2025)



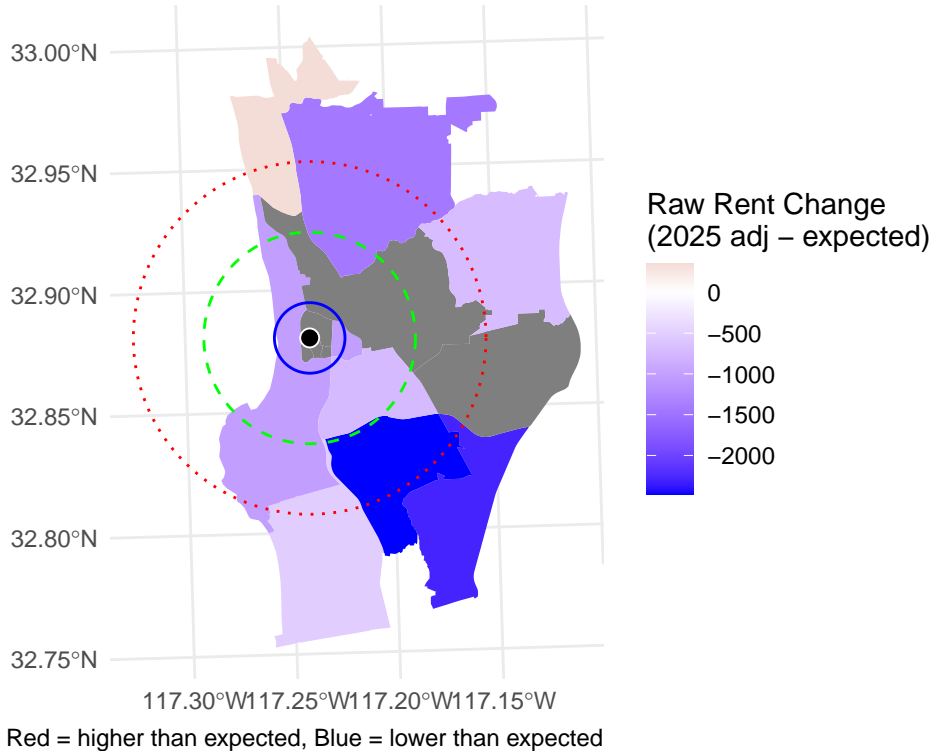
Real Rent Growth vs Expected near UC Merced
Inflation-adjusted (5% annual benchmark, 2015...2025)



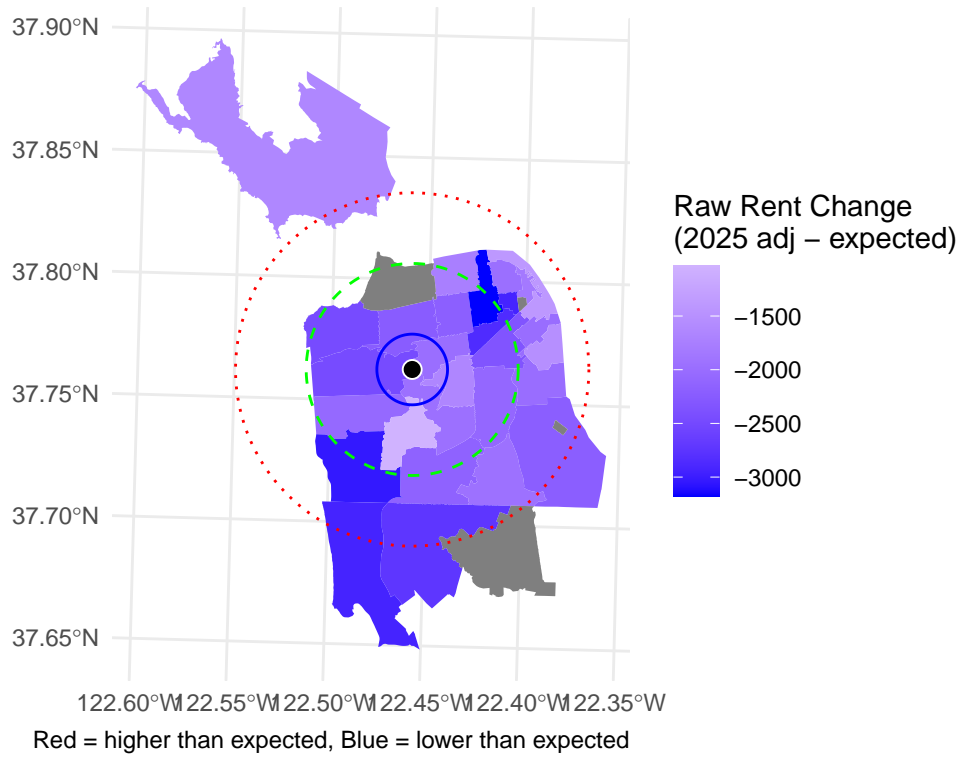
Real Rent Growth vs Expected near UC Riverside
Inflation-adjusted (5% annual benchmark, 2015...2025)



Real Rent Growth vs Expected near UC San Diego
Inflation-adjusted (5% annual benchmark, 2015...2025)

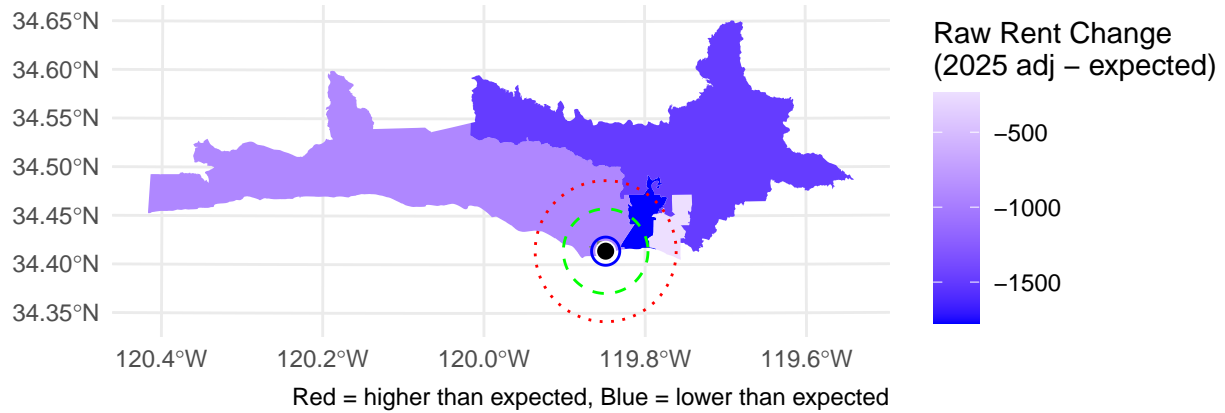


Real Rent Growth vs Expected near UC San Francisco Inflation-adjusted (5% annual benchmark, 2015...2025)



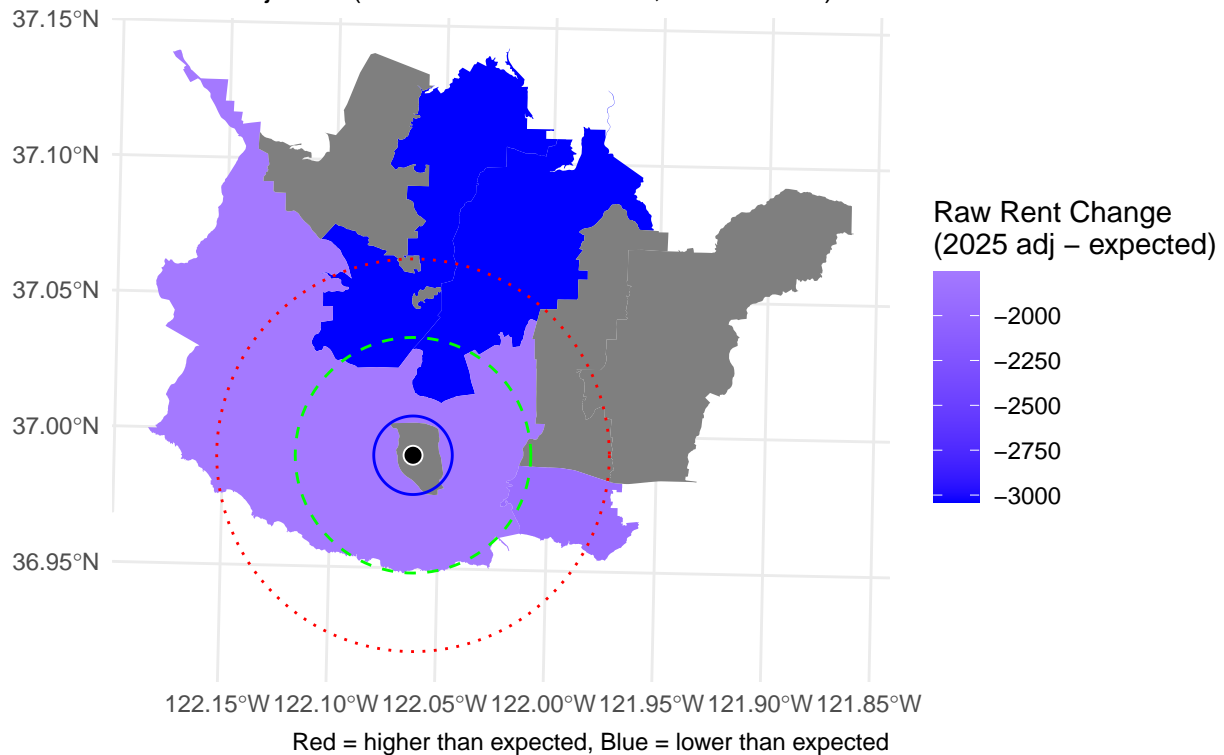
Real Rent Growth vs Expected near UC Santa Barbara

Inflation-adjusted (5% annual benchmark, 2015...2025)



Real Rent Growth vs Expected near UC Santa Cruz

Inflation-adjusted (5% annual benchmark, 2015...2025)



adjusted for inflation, rent increase, with percentage changes

```
options(tigris_use_cache = TRUE)
rent_data <- read_csv("clean_california_zillow_rent.csv") %>%
  filter(!is.na(RentIndex)) %>%
  mutate(ZIPCode = str_pad(as.character(ZIPCode), 5, pad = "0"))

# Step 1: Separate 2015 and 2025 data
rent_2015 <- rent_data %>%
  filter(Date == "2015-01-31") %>%
  select(ZIPCode, RentIndex_2015 = RentIndex)

rent_2025 <- rent_data %>%
  filter(Date == "2025-01-31") %>%
  select(ZIPCode, RentIndex_2025 = RentIndex)

# Step 2: Join the datasets by ZIP code
rent_joined <- inner_join(rent_2015, rent_2025, by = "ZIPCode") %>%
  mutate(
    RentIndex_2025_adj = RentIndex_2025 / (1.05 ^ 10),
    Expected_2025 = RentIndex_2015 * (1.05 ^ 10),
```

```

    Diff_vs_expected = RentIndex_2025 - Expected_2025
  )
zip_shapes <- zctas(cb = FALSE, starts_with = "9") %>%
  st_transform(3310) %>% st_simplify(dTolerance = 100)

# Check ZIP format
zip_shapes$ZCTA5CE20 <- as.character(zip_shapes$ZCTA5CE20)

common_zips <- intersect(unique(zip_shapes$ZCTA5CE20), unique(rent_joined$ZIPCode))
length(common_zips)

## [1] 898

head(common_zips)

## [1] "94587" "95126" "90602" "95130" "95035" "92657"

# Step 3: Adjust 2025 rent for inflation (CPI from BLS or estimate)
# https://www.bls.gov/regions/mid-atlantic/data/consumerpriceindexhistorical_us_table.htm
cpi_2015 <- 233.707
cpi_2025 <- 317.67
inflation_adjustment <- 1 + ((cpi_2025 - cpi_2015) / cpi_2015)

# Step 4: Compute expected rent under 5% annual increase
rent_joined <- rent_joined %>%
  mutate(
    Expected_2025 = RentIndex_2015 * (1 + 0.05)^10 * inflation_adjustment,
    Diff_vs_expected = RentIndex_2025 - Expected_2025,
    PercentDiff_vs_expected = 100 * (Diff_vs_expected) / Expected_2025
  )

# Step 5: Join with spatial ZIPs
zip_shapes$ZCTA5CE20 <- as.character(zip_shapes$ZCTA5CE20)
rent_joined$ZIPCode <- as.character(rent_joined$ZIPCode)

# Join with ZIP shapefile
rent_diff_sf <- left_join(zip_shapes, rent_joined, by = c("ZCTA5CE20" = "ZIPCode"))

for (i in 1:nrow(uc_sf)) {
  campus_name <- uc_sf$Campus[i]
  campus_pt <- uc_sf[i, ]

  buffer_1 <- st_buffer(campus_pt, 1609)
  buffer_3 <- st_buffer(campus_pt, 1609 * 3)
  buffer_5 <- st_buffer(campus_pt, 1609 * 9)

  # Filter nearby ZIPs for joined inflation-adjusted data
  zip_nearby <- rent_diff_sf[st_intersects(rent_diff_sf, buffer_5, sparse = FALSE), ]

  # Plot actual vs expected difference
  p <- ggplot() +

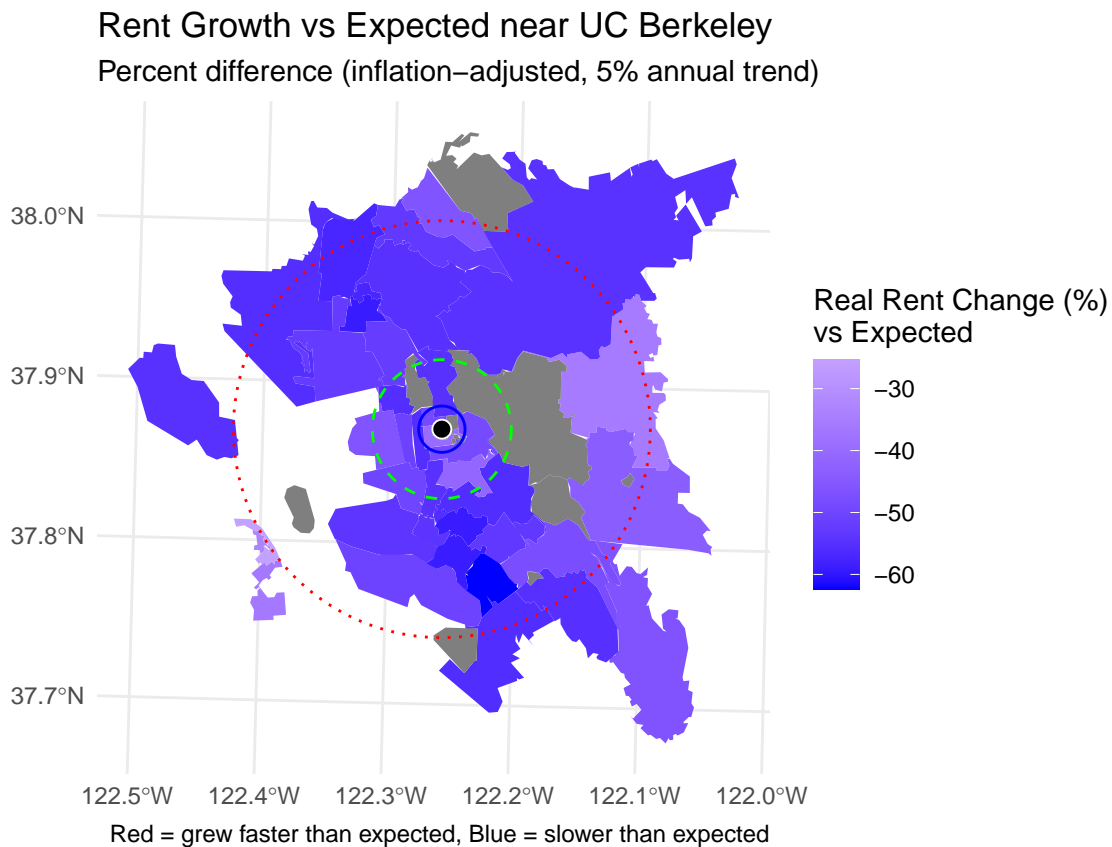
```

```

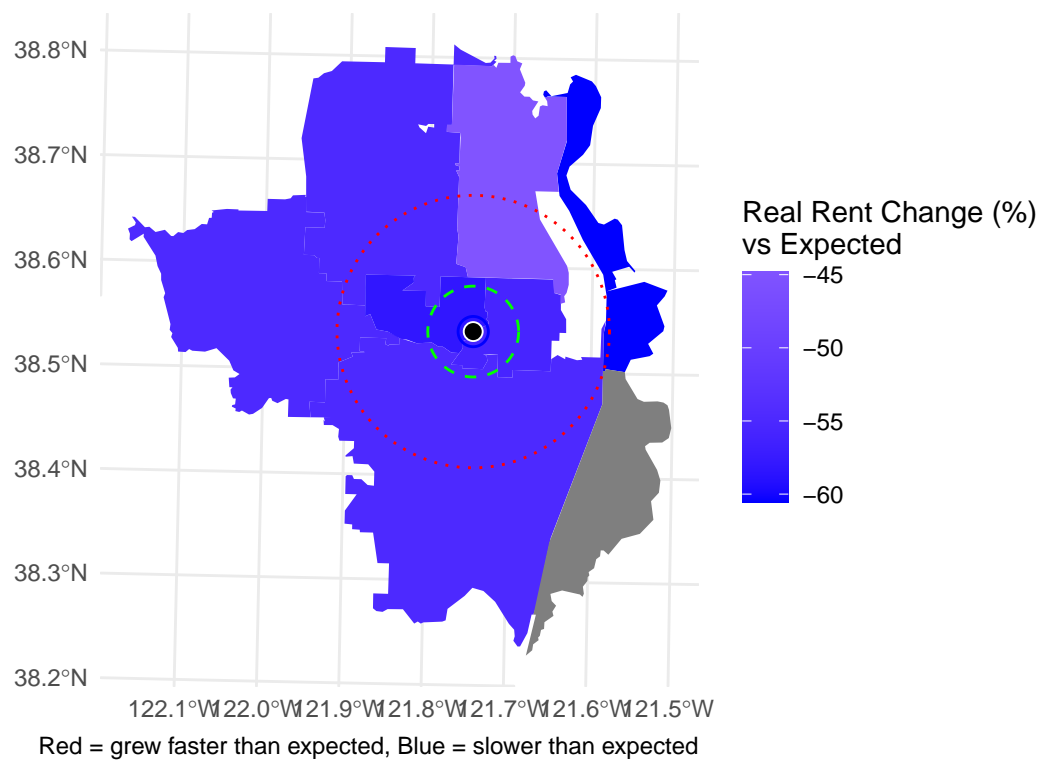
geom_sf(data = zip_nearby, aes(fill = PercentDiff_vs_expected), color = NA) +
scale_fill_gradient2(
  low = "blue", mid = "white", high = "darkred",
  midpoint = 0,
  name = "Real Rent Change (%) \nvs Expected"
) +
geom_sf(data = buffer_5, fill = NA, color = "red", linetype = "dotted", linewidth = 0.5) +
geom_sf(data = buffer_3, fill = NA, color = "green", linetype = "dashed", linewidth = 0.5) +
geom_sf(data = buffer_1, fill = NA, color = "blue", linetype = "solid", linewidth = 0.5) +
geom_sf(data = campus_pt, shape = 21, fill = "black", color = "white", size = 3) +
labs(
  title = paste0("Rent Growth vs Expected near ", campus_name),
  subtitle = "Percent difference (inflation-adjusted, 5% annual trend)",
  caption = "Red = grew faster than expected, Blue = slower than expected"
) +
theme_minimal()

print(p)
}

```

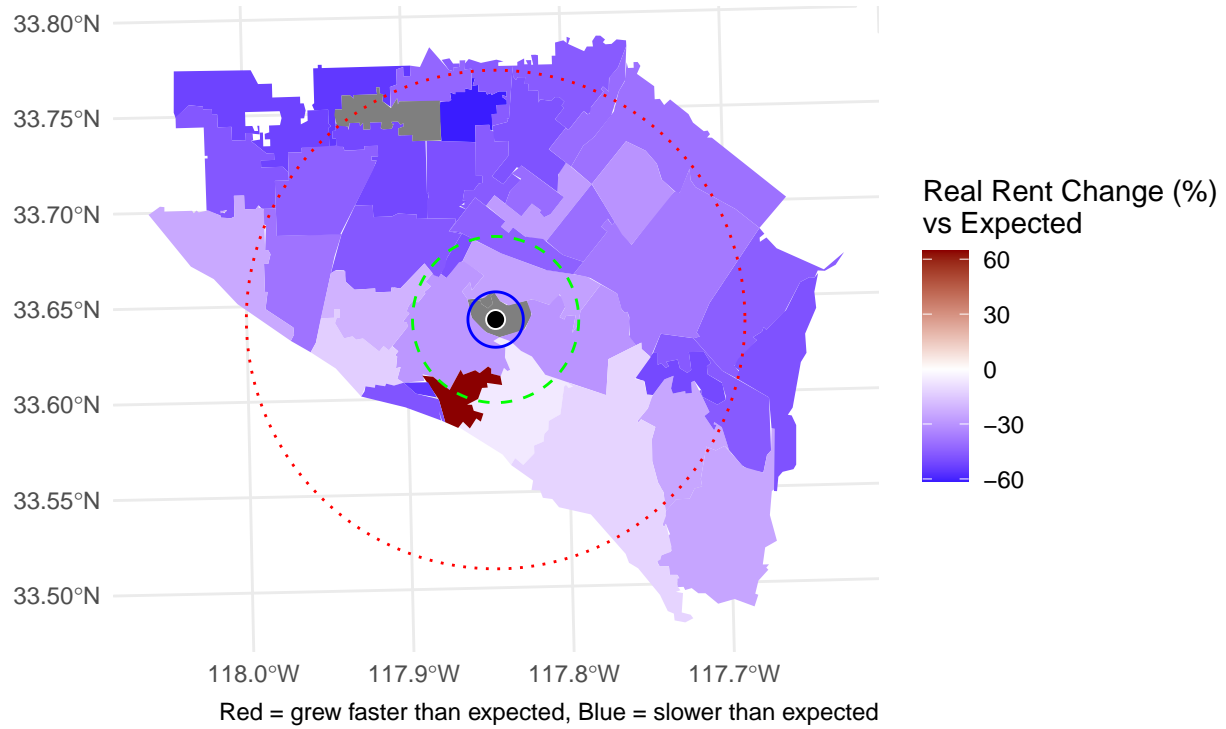


Rent Growth vs Expected near UC Davis
Percent difference (inflation-adjusted, 5% annual trend)



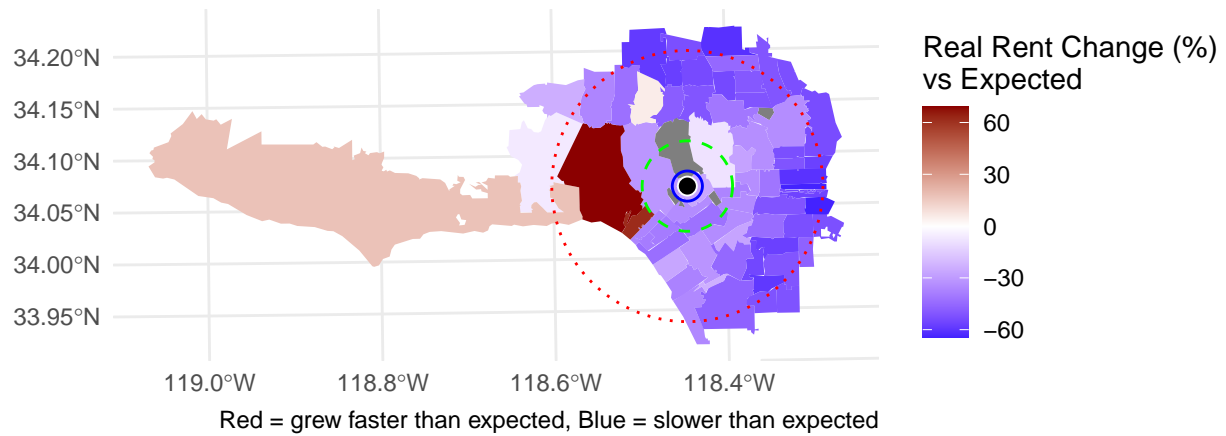
Rent Growth vs Expected near UC Irvine

Percent difference (inflation-adjusted, 5% annual trend)

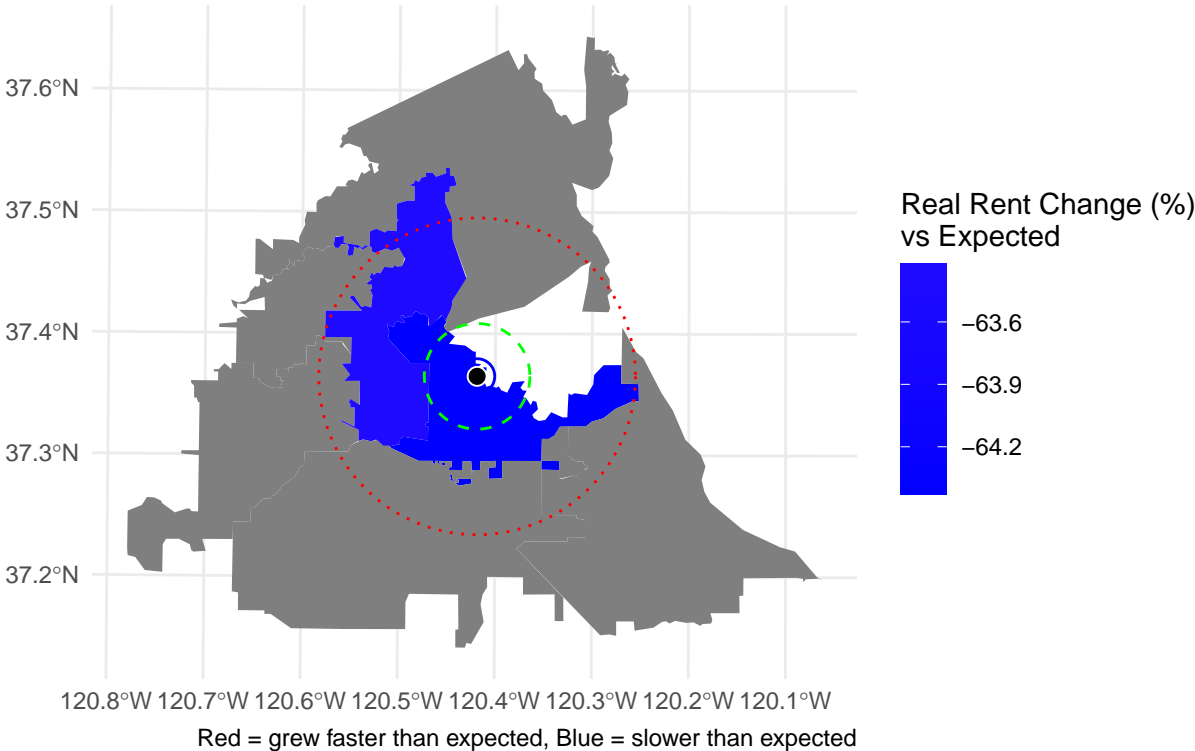


Rent Growth vs Expected near UCLA

Percent difference (inflation-adjusted, 5% annual trend)

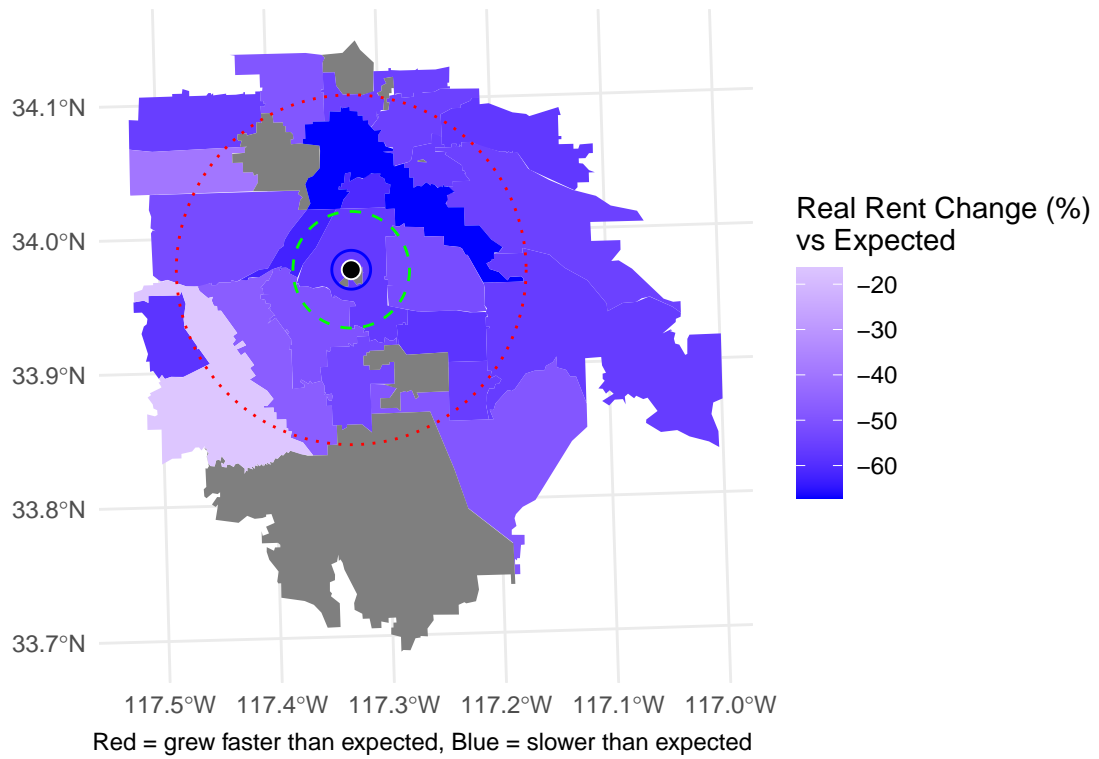


Rent Growth vs Expected near UC Merced
Percent difference (inflation-adjusted, 5% annual trend)



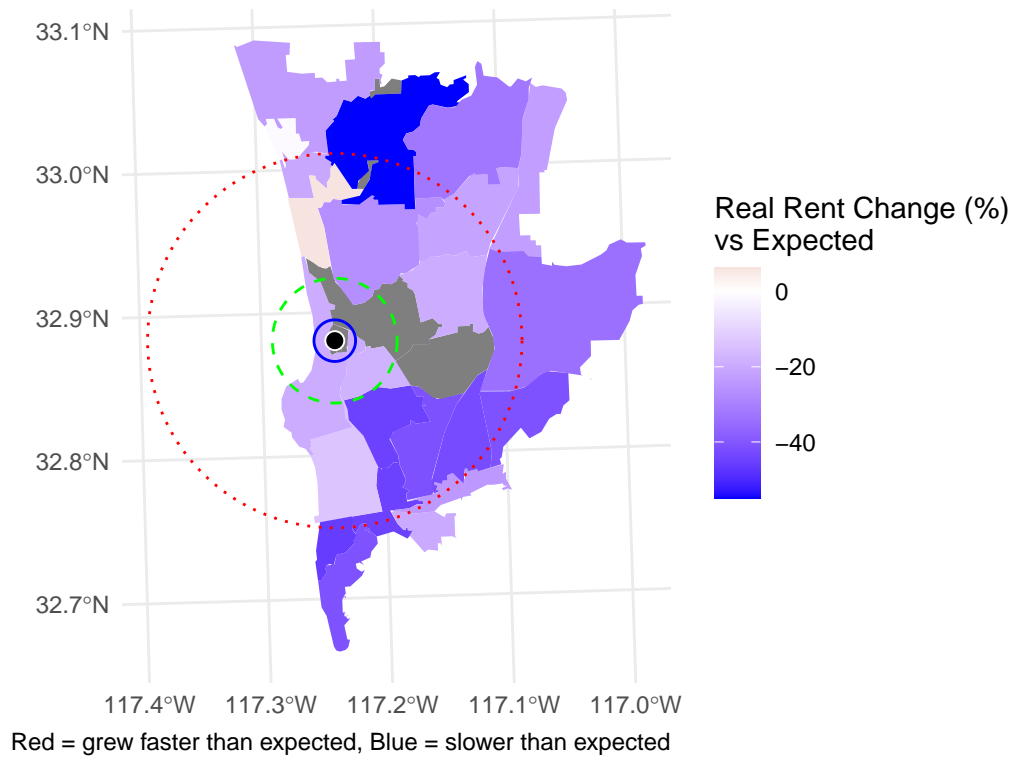
Rent Growth vs Expected near UC Riverside

Percent difference (inflation-adjusted, 5% annual trend)



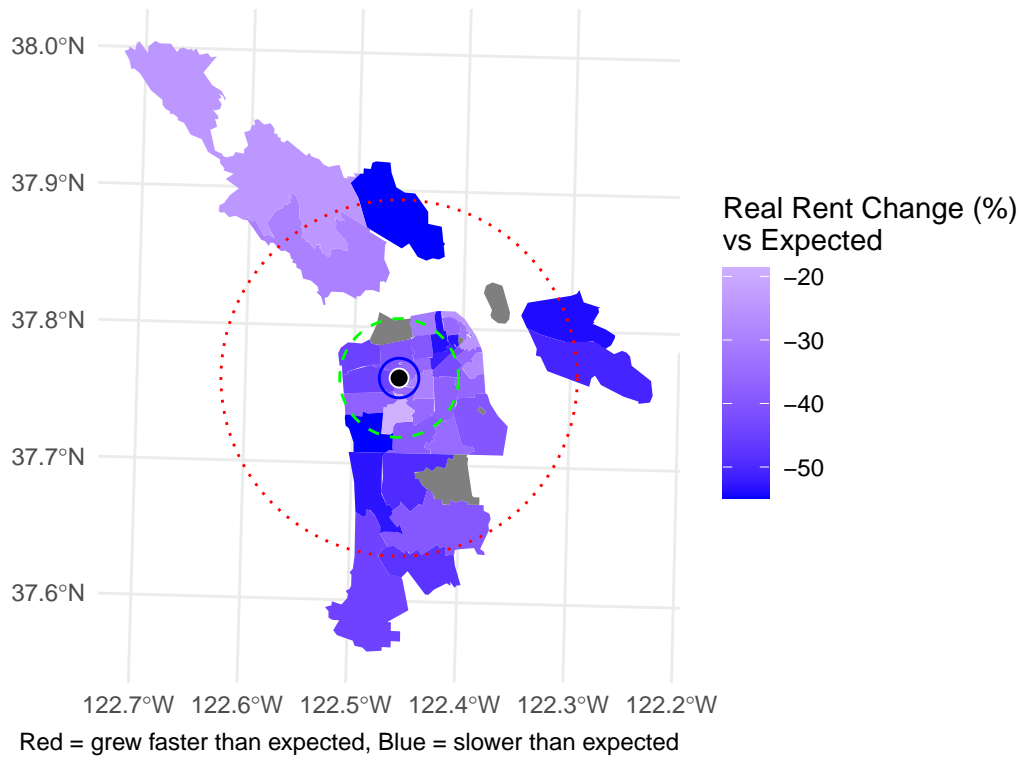
Rent Growth vs Expected near UC San Diego

Percent difference (inflation-adjusted, 5% annual trend)



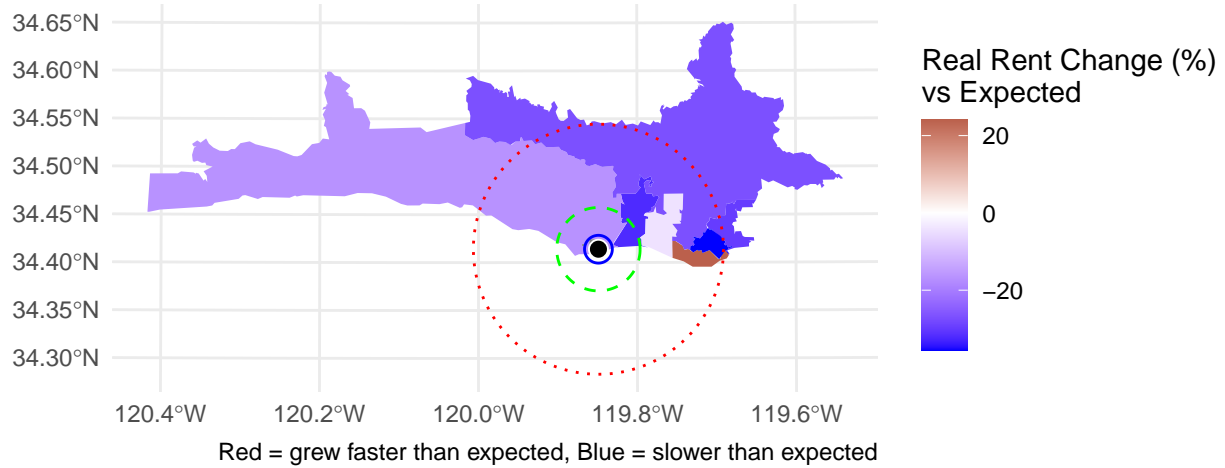
Rent Growth vs Expected near UC San Francisco

Percent difference (inflation-adjusted, 5% annual trend)



Rent Growth vs Expected near UC Santa Barbara

Percent difference (inflation-adjusted, 5% annual trend)



Rent Growth vs Expected near UC Santa Cruz

Percent difference (inflation-adjusted, 5% annual trend)

