



دانشگاه شهید بهشتی
Shahid Beheshti University

بررسی و تحلیل سیستم‌های انرژی تجدیدپذیر

برنامه‌سازی پیشرفته

گردآورندگان:

پونه رحیمی نژاد

محمد بابائی

مهدیس شاهکلایی

فهرست مطالب

۱.	مقدمه	۳
۱-۱	موضوع اصلی	۳
۱-۲	اهداف اصلی تحلیل	۳
۱-۳	آشنایی با متغیرها	۴
۲.	تحلیل توصیفی داده‌ها	۷
۲-۱	تحلیل عددی	۷
۲-۲	تحلیل بصری	۹
۳.	پیش‌پردازش داده‌ها	۱۷
۳-۱	پاکسازی داده‌ها	۱۷
۳-۲	مقادیر گم شده	۱۸
۳-۳	استانداردسازی داده‌ها	۱۸
۳-۴	تبدیل متغیرها	۱۹
۴.	الگوهای پنهان و همبستگی‌ها	۲۰
۴-۱	ماتریس همبستگی	۲۰
۴-۲	نمودار پراکندگی	۲۱
۴-۳	تست‌های آماری	۲۲
۵.	بصری سازی نتایج	۲۵
۵-۱	نمودار خطی	۲۵
۵-۲	نمودار ویولنی	۳۱
۵-۳	نمودار شبکه‌ای	۳۲
۵-۴	نمودار دایره‌ای	۳۴
۶.	کاربرد نتایج در تصمیم‌گیری	۳۴
۷.	محدو دیت‌های مجموعه داده	۳۶

۱. مقدمه

تحلیل داده‌های مربوط به منابع انرژی تجدیدپذیر از اهمیت ویژه‌ای برخوردار است، زیرا این منابع نقش حیاتی در تامین انرژی پاک و کاهش آلودگی‌های زیستمحیطی دارند. در این پژوهه، هدف اصلی، بررسی و تحلیل داده‌های مربوط به انواع مختلف انرژی تجدیدپذیر از جمله خورشیدی، بادی، هیدروالکتریک، زمین‌گرمایی، زیستتوده، جزر و مد و موج است. داده‌های موجود شامل اطلاعاتی در مورد ظرفیت نصب شده، تولید و مصرف سالیانه انرژی، ظرفیت ذخیره‌سازی انرژی، کارایی سیستم‌های ذخیره‌سازی، سطح یکپارچگی شبکه، هزینه‌های سرمایه‌گذاری اولیه، منابع تامین مالی، مشوق‌های مالی، کاهش انتشار گازهای گلخانه‌ای، شاخص کاهش آلودگی هوا و تعداد مشاغل ایجاد شده می‌باشد.

۱-۱ موضوع اصلی

موضوع اصلی این تحلیل بررسی داده‌های جامع مربوط به منابع انرژی تجدیدپذیر و استخراج الگوها و همبستگی‌های مهم در میان متغیرهای مختلف است. این تحلیل به ما کمک می‌کند تا درک بهتری از عملکرد و کارایی هر یک از این منابع به دست آوریم و بتوانیم نقاط قوت و ضعف موجود را شناسایی کنیم.

۲-۱ اهداف اصلی تحلیل

- شناسایی چالش‌ها و اهداف: درک چالش‌های خاص مربوط به تحلیل داده‌های انرژی تجدیدپذیر و تعیین اهداف پژوهشی مشخص. این مرحله شامل شناسایی سوالات کلیدی است که قصد داریم به آنها پاسخ دهیم و همچنین تعیین معیارهای موفقیت برای تحلیل.
- شناسایی متغیرهای کلیدی و توزیع آنها: شناسایی متغیرهای کلیدی موجود در داده‌ها و بررسی توزیع آنها. این شامل تحلیل توزیع‌های آماری و بررسی نحوه تعامل این متغیرها با یکدیگر می‌شود. با این کار، می‌توانیم الگوهای مهم و ارتباطات بین متغیرها را شناسایی کنیم.
- تحلیل داده‌های توصیفی: انجام تحلیل‌های توصیفی برای درک طبیعت داده‌ها و الگوهای داخلی آنها. این تحلیل‌ها شامل محاسبه مقادیر مرکزی مانند میانگین، میانه و مد، شناسایی ناهنجاری‌ها و بررسی همبستگی‌های بین

متغیرها می‌باشد. هدف این است که بتوانیم به تصویری جامع از داده‌ها دست یابیم و بینش‌های ارزشمندی استخراج کنیم.

این تحلیل به ما کمک می‌کند تا به درک عمیق‌تری از داده‌های انرژی تجدیدپذیر برسیم و بتوانیم از این بینش‌ها برای تصمیم‌گیری‌های عملی و بهبود سیاست‌های مربوط به انرژی تجدیدپذیر استفاده کنیم. همچنین، با شناسایی نقاط قوت و ضعف موجود در سیستم‌های انرژی تجدیدپذیر، می‌توانیم راهکارهای عملی برای بهبود کارایی و کاهش آلودگی‌های زیست‌محیطی ارائه دهیم. نتایج این تحلیل می‌تواند به سیاست‌گذاران و مدیران کمک کند تا تصمیمات بهتری در زمینه توسعه و بهره‌برداری از منابع انرژی تجدیدپذیر اتخاذ کنند و به دستیابی به اهداف پایداری زیست‌محیطی کمک کنند.

با انجام این پروژه، امیدواریم بتوانیم به درک بهتری از عملکرد و کارایی منابع انرژی تجدیدپذیر برسیم و نقش موثری در بهبود وضعیت انرژی و محیط زیست ایفا کنیم. این تحلیل نه تنها به شناسایی چالش‌های موجود کمک می‌کند، بلکه می‌تواند راهکارهای نوآورانه‌ای برای مواجهه با این چالش‌ها ارائه دهد و به توسعه پایدار انرژی کمک کند.

۳-۱ آشنایی با متغیرها

۱. Type_of_Renewable_Energy:

این متغیر نشان‌دهنده نوع انرژی تجدیدپذیر مورد استفاده در پروژه است که متغیری رسته‌ای و دارای ۷ سطح است:

۱) انرژی خورشیدی ۲) انرژی بادی ۳) انرژی هیدرولکتریک ۴) انرژی زمین گرمایی ۵) انرژی زیست‌توده ۶) انرژی جزر و مدی ۷) انرژی موج

۲. Installed_Capacity_MW:

این متغیر میزان ظرفیت نصب شده پروژه را به مگاوات (MW) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۳. Energy_Production_MWh:

این متغیر میزان تولید انرژی را به مگاوات در ساعت (MWh) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۴. Energy_Consumption_MWh:

این متغیر میزان مصرف انرژی را به مگاوات در ساعت (MWh) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۵. Energy_Storage_Capacity_MWh:

این متغیر ظرفیت ذخیره‌سازی انرژی را به مگاوات ساعت (MWh) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۶. Storage_Efficiency_Percentage:

این متغیر درصد بازدهی سیستم ذخیره‌سازی انرژی را نشان می‌دهد و از نوع عددی پیوسته (float 64) است. بازدهی ذخیره‌سازی انرژی نشان‌دهنده کارایی سیستم در تبدیل و ذخیره انرژی است و به تحلیل عملکرد سیستم کمک می‌کند.

۷. Grid_Integration_Level:

این متغیر سطح یکپارچگی با شبکه برق را نشان می‌دهد و متغیری رسته‌ای شامل ۴ سطح مختلف است:

- (۱) یکپارچگی کامل (کاملاً ادغام شده)
- (۲) یکپارچگی جزئی (نیمه ادغام شده)
- (۳) یکپارچگی کم (کمینه ادغام شده)
- (۴) شبکه جداگانه

سطح یکپارچگی با شبکه برق به میزان و کیفیت اتصال و ادغام منابع انرژی، خصوصاً منابع تجدیدپذیر، به شبکه برق اشاره دارد. این سطح نشان‌دهنده توانایی شبکه برق در مدیریت، کنترل و بهره‌برداری از این منابع به‌گونه‌ای است که پایداری و بهره‌وری سیستم حفظ شود.

۸. Initial_Investment_USD:

این متغیر میزان سرمایه‌گذاری اولیه را به دلار آمریکا (USD) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۹. Funding_Sources:

این متغیر رسته‌ای، منابع مالی را نشان می‌دهد و شامل ۳ دسته مختلف است: ۱) دولتی ۲) خصوصی ۳) ترکیبی عمومی و خصوصی (مانند وام و ...)

۱۰. Financial_Incentives_USD:

این متغیر میزان مشوق‌های مالی ارائه شده را به دلار آمریکا (USD) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۱۱. GHG_Emission_Reduction_tCO2e:

این متغیر میزان کاهش انتشار گازهای گلخانه‌ای را به تن CO₂ معادل (tCO₂e) نشان می‌دهد و از نوع عددی پیوسته (float 64) است.

۱۲. Air_Pollution_Reduction_Index:

این متغیر شاخص کاهش آلودگی هوا را نشان می‌دهد و از نوع عددی پیوسته (float 64) است که بین ۱ تا ۱۰۰ است.

۱۳. Jobs_Created:

این متغیر تعداد شغل‌های ایجاد شده توسط پروژه را نشان می‌دهد و از نوع عددی گسسته (int 64) است.

۲. تحلیل توصیفی داده‌ها

۱-۲ تحلیل عددی

پس از فراخوانی داده‌ها و بارگذاری به تحلیل آن‌ها می‌پردازیم. تحلیل عددی یکی از مهمترین مراحل در تحلیل داده‌ها است که به ما کمک می‌کند تا دید عمیق‌تری نسبت به توزیع و رفتار داده‌ها داشته باشیم. در اینجا، ما از معیارهای آماری مختلفی مانند میانگین، میانه، مد، دامنه، واریانس و انحراف معیار استفاده می‌کنیم تا توزیع و گرایش مرکزی داده‌ها را بهتر بفهمیم. بعد از چک کردن نوع داده‌ها، داده‌های گسسته (Type of Renewable Energy, Grid Integration) به تغییر می‌کنند تا در روند تحلیل اختلالی ایجاد نشود.

Type_of_Renewable_Energy	int64
Installed_Capacity_MW	float64
Energy_Production_MWh	float64
Energy_Consumption_MWh	float64
Energy_Storage_Capacity_MWh	float64
Storage_Efficiency_Percentage	float64
Grid_Integration_Level	int64
Initial_Investment_USD	float64
Funding_Sources	int64
Financial_Incentives_USD	float64
GHG_Emission_Reduction_tCO2e	float64
Air_Pollution_Reduction_Index	float64
Jobs_Created	int64
dtype: object	



Type_of_Renewable_Energy	category
Installed_Capacity_MW	float64
Energy_Production_MWh	float64
Energy_Consumption_MWh	float64
Energy_Storage_Capacity_MWh	float64
Storage_Efficiency_Percentage	float64
Grid_Integration_Level	category
Initial_Investment_USD	float64
Funding_Sources	category
Financial_Incentives_USD	float64
GHG_Emission_Reduction_tCO2e	float64
Air_Pollution_Reduction_Index	float64
Jobs_Created	int64
dtype: object	

<p>Median:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td>Installed_Capacity_MW</td> <td>492.91</td> </tr> <tr> <td>Energy_Production_MWh</td> <td>253216.87</td> </tr> <tr> <td>Energy_Consumption_MWh</td> <td>225226.37</td> </tr> <tr> <td>Energy_Storage_Capacity_MWh</td> <td>5054.04</td> </tr> <tr> <td>Storage_Efficiency_Percentage</td> <td>75.28</td> </tr> <tr> <td>Initial_Investment_USD</td> <td>253990954.45</td> </tr> <tr> <td>Financial_Incentives_USD</td> <td>10024143.54</td> </tr> <tr> <td>GHG_Emission_Reduction_tCO2e</td> <td>25424.48</td> </tr> <tr> <td>Air_Pollution_Reduction_Index</td> <td>50.25</td> </tr> <tr> <td>Jobs_Created</td> <td>2496.00</td> </tr> </tbody> </table> <p>dtype: float64</p>	Installed_Capacity_MW	492.91	Energy_Production_MWh	253216.87	Energy_Consumption_MWh	225226.37	Energy_Storage_Capacity_MWh	5054.04	Storage_Efficiency_Percentage	75.28	Initial_Investment_USD	253990954.45	Financial_Incentives_USD	10024143.54	GHG_Emission_Reduction_tCO2e	25424.48	Air_Pollution_Reduction_Index	50.25	Jobs_Created	2496.00	<p>Mode:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td>Type_of_Renewable_Energy</td> <td>2.00</td> </tr> <tr> <td>Installed_Capacity_MW</td> <td>1.09</td> </tr> <tr> <td>Energy_Production_MWh</td> <td>1030.10</td> </tr> <tr> <td>Energy_Consumption_MWh</td> <td>584.05</td> </tr> <tr> <td>Energy_Storage_Capacity_MWh</td> <td>2.20</td> </tr> <tr> <td>Storage_Efficiency_Percentage</td> <td>50.00</td> </tr> <tr> <td>Grid_Integration_Level</td> <td>1.00</td> </tr> <tr> <td>Initial_Investment_USD</td> <td>1008107.10</td> </tr> <tr> <td>Funding_Sources</td> <td>3.00</td> </tr> <tr> <td>Financial_Incentives_USD</td> <td>51558.42</td> </tr> <tr> <td>GHG_Emission_Reduction_tCO2e</td> <td>100.97</td> </tr> <tr> <td>Air_Pollution_Reduction_Index</td> <td>1.01</td> </tr> <tr> <td>Jobs_Created</td> <td>946.00</td> </tr> </tbody> </table> <p>Name: 0, dtype: float64</p>	Type_of_Renewable_Energy	2.00	Installed_Capacity_MW	1.09	Energy_Production_MWh	1030.10	Energy_Consumption_MWh	584.05	Energy_Storage_Capacity_MWh	2.20	Storage_Efficiency_Percentage	50.00	Grid_Integration_Level	1.00	Initial_Investment_USD	1008107.10	Funding_Sources	3.00	Financial_Incentives_USD	51558.42	GHG_Emission_Reduction_tCO2e	100.97	Air_Pollution_Reduction_Index	1.01	Jobs_Created	946.00	<p>Mean:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td>Installed_Capacity_MW</td> <td>495.86</td> </tr> <tr> <td>Energy_Production_MWh</td> <td>252350.96</td> </tr> <tr> <td>Energy_Consumption_MWh</td> <td>225981.85</td> </tr> <tr> <td>Energy_Storage_Capacity_MWh</td> <td>5030.20</td> </tr> <tr> <td>Storage_Efficiency_Percentage</td> <td>75.22</td> </tr> <tr> <td>Initial_Investment_USD</td> <td>251484914.81</td> </tr> <tr> <td>Financial_Incentives_USD</td> <td>10029766.26</td> </tr> <tr> <td>GHG_Emission_Reduction_tCO2e</td> <td>25234.72</td> </tr> <tr> <td>Air_Pollution_Reduction_Index</td> <td>50.72</td> </tr> <tr> <td>Jobs_Created</td> <td>2502.67</td> </tr> </tbody> </table> <p>dtype: float64</p>	Installed_Capacity_MW	495.86	Energy_Production_MWh	252350.96	Energy_Consumption_MWh	225981.85	Energy_Storage_Capacity_MWh	5030.20	Storage_Efficiency_Percentage	75.22	Initial_Investment_USD	251484914.81	Financial_Incentives_USD	10029766.26	GHG_Emission_Reduction_tCO2e	25234.72	Air_Pollution_Reduction_Index	50.72	Jobs_Created	2502.67	<p>Variance:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td>Installed_Capacity_MW</td> <td>83066.66</td> </tr> <tr> <td>Energy_Production_MWh</td> <td>20754123600.75</td> </tr> <tr> <td>Energy_Consumption_MWh</td> <td>16698758955.20</td> </tr> <tr> <td>Energy_Storage_Capacity_MWh</td> <td>8375793.55</td> </tr> <tr> <td>Storage_Efficiency_Percentage</td> <td>209.82</td> </tr> <tr> <td>Initial_Investment_USD</td> <td>20530382616723924.00</td> </tr> <tr> <td>Financial_Incentives_USD</td> <td>33492871151090.72</td> </tr> <tr> <td>GHG_Emission_Reduction_tCO2e</td> <td>206753204.55</td> </tr> <tr> <td>Air_Pollution_Reduction_Index</td> <td>815.48</td> </tr> <tr> <td>Jobs_Created</td> <td>2106018.19</td> </tr> </tbody> </table> <p>dtype: float64</p>	Installed_Capacity_MW	83066.66	Energy_Production_MWh	20754123600.75	Energy_Consumption_MWh	16698758955.20	Energy_Storage_Capacity_MWh	8375793.55	Storage_Efficiency_Percentage	209.82	Initial_Investment_USD	20530382616723924.00	Financial_Incentives_USD	33492871151090.72	GHG_Emission_Reduction_tCO2e	206753204.55	Air_Pollution_Reduction_Index	815.48	Jobs_Created	2106018.19	<p>Range:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td>Installed_Capacity_MW</td> <td>998.89</td> </tr> <tr> <td>Energy_Production_MWh</td> <td>498961.10</td> </tr> <tr> <td>Energy_Consumption_MWh</td> <td>449338.62</td> </tr> <tr> <td>Energy_Storage_Capacity_MWh</td> <td>9996.94</td> </tr> <tr> <td>Storage_Efficiency_Percentage</td> <td>49.99</td> </tr> <tr> <td>Initial_Investment_USD</td> <td>498932604.50</td> </tr> <tr> <td>Financial_Incentives_USD</td> <td>19946994.61</td> </tr> <tr> <td>GHG_Emission_Reduction_tCO2e</td> <td>49896.60</td> </tr> <tr> <td>Air_Pollution_Reduction_Index</td> <td>98.97</td> </tr> <tr> <td>Jobs_Created</td> <td>4989.00</td> </tr> </tbody> </table> <p>dtype: float64</p>	Installed_Capacity_MW	998.89	Energy_Production_MWh	498961.10	Energy_Consumption_MWh	449338.62	Energy_Storage_Capacity_MWh	9996.94	Storage_Efficiency_Percentage	49.99	Initial_Investment_USD	498932604.50	Financial_Incentives_USD	19946994.61	GHG_Emission_Reduction_tCO2e	49896.60	Air_Pollution_Reduction_Index	98.97	Jobs_Created	4989.00
Installed_Capacity_MW	492.91																																																																																																													
Energy_Production_MWh	253216.87																																																																																																													
Energy_Consumption_MWh	225226.37																																																																																																													
Energy_Storage_Capacity_MWh	5054.04																																																																																																													
Storage_Efficiency_Percentage	75.28																																																																																																													
Initial_Investment_USD	253990954.45																																																																																																													
Financial_Incentives_USD	10024143.54																																																																																																													
GHG_Emission_Reduction_tCO2e	25424.48																																																																																																													
Air_Pollution_Reduction_Index	50.25																																																																																																													
Jobs_Created	2496.00																																																																																																													
Type_of_Renewable_Energy	2.00																																																																																																													
Installed_Capacity_MW	1.09																																																																																																													
Energy_Production_MWh	1030.10																																																																																																													
Energy_Consumption_MWh	584.05																																																																																																													
Energy_Storage_Capacity_MWh	2.20																																																																																																													
Storage_Efficiency_Percentage	50.00																																																																																																													
Grid_Integration_Level	1.00																																																																																																													
Initial_Investment_USD	1008107.10																																																																																																													
Funding_Sources	3.00																																																																																																													
Financial_Incentives_USD	51558.42																																																																																																													
GHG_Emission_Reduction_tCO2e	100.97																																																																																																													
Air_Pollution_Reduction_Index	1.01																																																																																																													
Jobs_Created	946.00																																																																																																													
Installed_Capacity_MW	495.86																																																																																																													
Energy_Production_MWh	252350.96																																																																																																													
Energy_Consumption_MWh	225981.85																																																																																																													
Energy_Storage_Capacity_MWh	5030.20																																																																																																													
Storage_Efficiency_Percentage	75.22																																																																																																													
Initial_Investment_USD	251484914.81																																																																																																													
Financial_Incentives_USD	10029766.26																																																																																																													
GHG_Emission_Reduction_tCO2e	25234.72																																																																																																													
Air_Pollution_Reduction_Index	50.72																																																																																																													
Jobs_Created	2502.67																																																																																																													
Installed_Capacity_MW	83066.66																																																																																																													
Energy_Production_MWh	20754123600.75																																																																																																													
Energy_Consumption_MWh	16698758955.20																																																																																																													
Energy_Storage_Capacity_MWh	8375793.55																																																																																																													
Storage_Efficiency_Percentage	209.82																																																																																																													
Initial_Investment_USD	20530382616723924.00																																																																																																													
Financial_Incentives_USD	33492871151090.72																																																																																																													
GHG_Emission_Reduction_tCO2e	206753204.55																																																																																																													
Air_Pollution_Reduction_Index	815.48																																																																																																													
Jobs_Created	2106018.19																																																																																																													
Installed_Capacity_MW	998.89																																																																																																													
Energy_Production_MWh	498961.10																																																																																																													
Energy_Consumption_MWh	449338.62																																																																																																													
Energy_Storage_Capacity_MWh	9996.94																																																																																																													
Storage_Efficiency_Percentage	49.99																																																																																																													
Initial_Investment_USD	498932604.50																																																																																																													
Financial_Incentives_USD	19946994.61																																																																																																													
GHG_Emission_Reduction_tCO2e	49896.60																																																																																																													
Air_Pollution_Reduction_Index	98.97																																																																																																													
Jobs_Created	4989.00																																																																																																													

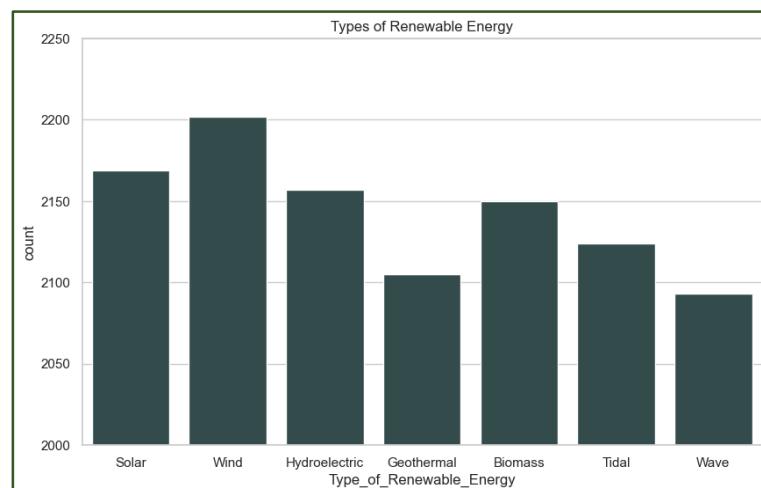
تحلیل عددی به ما کمک می‌کند تا بفهمیم توزیع داده‌ها چگونه است و آیا داده‌ها به طور کلی به یک سمت گرایش دارند یا خیر. با استفاده از این تحلیل، می‌توانیم تصمیمات بهتری در زمینه پیش‌پردازش داده‌ها، مدل‌سازی و تفسیر نتایج بگیریم. برای مثال، اگر واریانس یک متغیر بسیار بالا باشد، ممکن است نیاز به نرمال‌سازی یا استانداردسازی آن متغیر باشد تا مدل‌ها بتوانند بهتر عمل کنند. همچنین، با بررسی میانگین و میانه می‌توانیم بفهمیم که آیا داده‌ها دارای انحراف از نوع خاصی هستند یا خیر.

۱-۲ تحلیل بصری

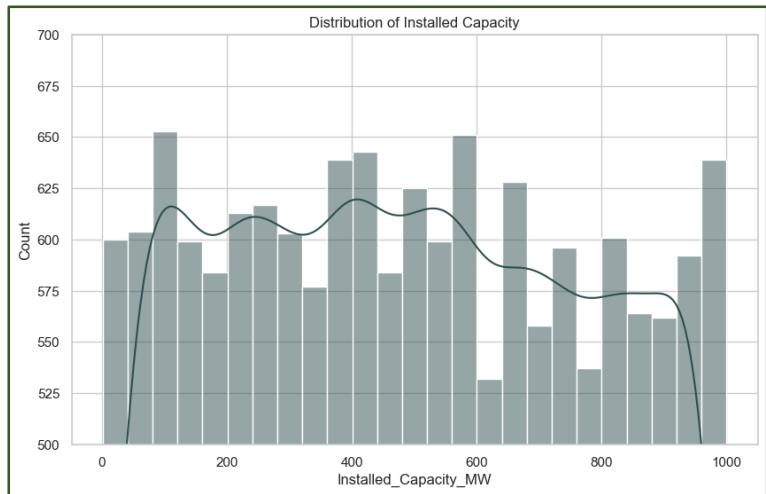
با مصورسازی داده‌ها می‌توان به شناسایی الگوهای شناسایی داده‌های پرت و ... پرداخت. نمودارها با جذابیت بیشتر و درک آسان‌تر به ما کمک می‌کنند تا متغیرها را تحلیل و بررسی کیم و تصمیمات بهتری بگیریم. در ادامه نتایج تحلیل‌های عددی و نمودارها بررسی شده‌اند.

Type of Renewable Energy (۱)

انرژی بادی رایج‌ترین نوع انرژی تجدیدپذیر در این مجموعه داده‌ها است. انرژی موجی کمترین میزان را دارد. دیگر انواع انرژی تجدیدپذیر (خورشیدی، برق‌آبی، زمین‌گرمایی، زیست‌توده، جزر و مدی) دارای تعداد پروژه‌هایی بین این دو مقدار



هستند که نشان‌دهنده بازه این متغیر می‌باشند. این نمودار تصویری واضح از توزیع پروژه‌های مختلف انرژی تجدیدپذیر ارائه می‌دهد و نشان می‌دهد که به طور کل از هر انرژی بین ۲۰۵۰ تا ۲۱۵۰ عدد در کل داده وجود دارد.



Capacity Installed (MW) (۲)

- میانگین: ۴۹۵,۸۶

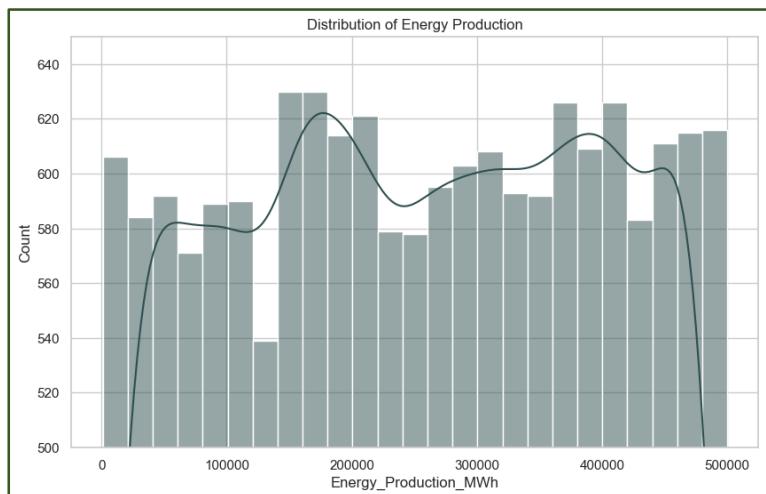
- میانه: ۴۹۲,۹۱

- انحراف معیار: ۲۸۸,۲۱

- بازه: ۹۹۸,۸۹

ظرفیت نصب شده در پروژه‌ها به طور میانگین ۴۹۵,۸۶ مگاوات است و میانه نزدیک به میانگین است که نشان می‌دهد توزیع داده‌ها نسبتاً متقارن است. با توجه به انحراف معیار و بازه گستردگی، تنوع زیادی در ظرفیت‌های نصب شده وجود دارد.

ظرفیت‌های نصب شده در محدوده ۰ تا ۱۰۰۰ MW پراکنده شده‌اند. توزیع ظرفیت‌ها به صورت تقریباً یکنواخت در بیشتر محدوده‌ها دیده می‌شود، با برخی پیک‌های قابل توجه در حدود ۱۰۰ MW، ۴۰۰ MW، و نزدیک ۱۰۰۰ MW بالاترین ظرفیت‌های نصب شده در حدود ۰ تا ۲۰۰ MW و نزدیک به ۱۰۰۰ MW دیده می‌شود. در میانه محدوده، بین ۶۰۰ تا ۸۰۰ MW، چگالی پایین‌تری مشاهده می‌شود.



Energy Production (MWh) (۳)

- میانگین: ۲۵۲۳۵۰,۹۶

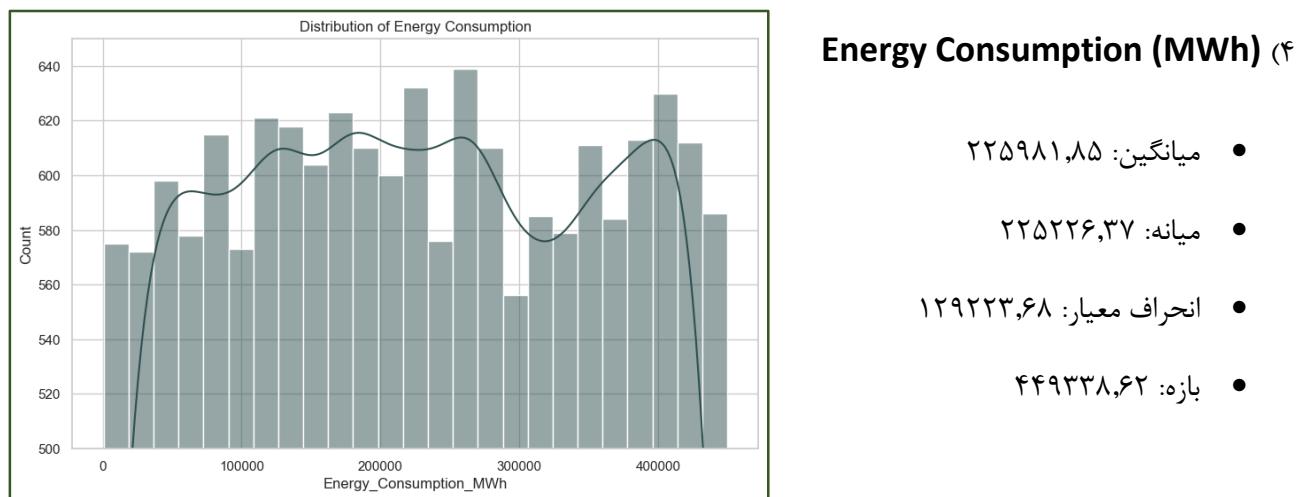
- میانه: ۲۵۳۲۱۶,۸۷

- انحراف معیار: ۱۴۴۰۶۲,۹۲

- بازه: ۴۹۸۹۶۱,۱۰

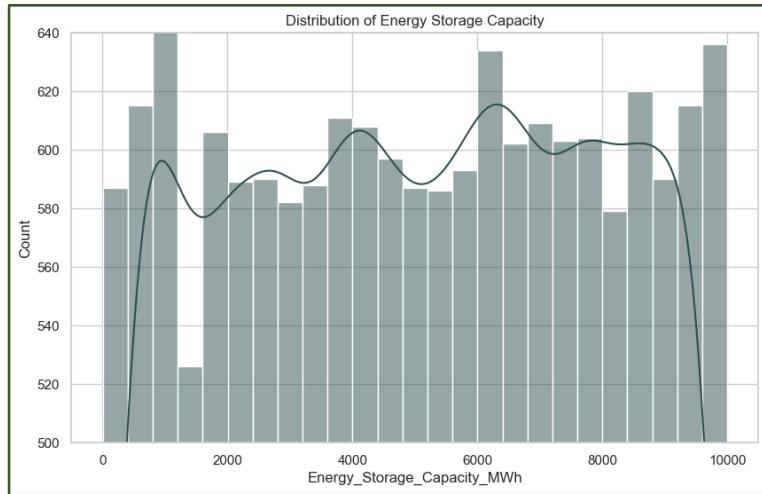
تولید انرژی به طور میانگین ۲۵۲۳۵۰,۹۶ مگاوات در ساعت است و میانه نیز تقریباً با میانگین برابر است. این امر نشان‌دهنده توزیع نسبتاً متقارن داده‌هاست. بازه و انحراف معیار بالا نشان‌دهنده تفاوت‌های قابل توجه در میزان تولید انرژی پروژه‌هاست.

پراکندگی نمودار تولیدات انرژی در محدوده ۰ الی ۵۰۰۰۰۰ MWh می‌باشد. واریانس نمودار بسیار بالا بوده و عدم همگنی بین پروژه‌ها و تولیدات انرژی وجود دارد و به همین سبب الگو ارتباطی وجود نداشته و داده‌های این متغیر قابل پیش‌بینی نخواهند بود. پیک قابل توجه در حدود ۱۹۰۰۰۰ MWh مشاهده می‌شود و بعد پیک دوم در میانگین بازه پیش‌بینی نخواهد بود. دیده می‌شود که بطور قابل توجهی از میانه و میانگین بیشتر هستند.



صرف انرژی در پروژه‌ها به طور میانگین ۲۲۵۹۸۱,۸۵ مگاوات در ساعت است. میانه نیز نزدیک به میانگین است که نشان‌دهنده توزیع نسبتاً متقارن داده‌هاست. با توجه به انحراف معیار و بازه گسترده، تفاوت‌های زیادی در صرف انرژی پروژه‌ها مشاهده می‌شود.

پراکندگی نمودار صرف انرژی در محدوده ۰-۵۰۰۰۰۰ MWh می‌باشد و با توجه به واریانس بالای صرف انرژی، پروژه‌ها در سطوح متفاوتی انرژی صرف کرده‌اند، مینیمم این نمودار در محدوده ۳۰۰۰۰۰ الی ۳۴۰۰۰۰ MWh بوده. در باقی سطوح صرف می‌باشد که به این معناست که به نسبت پروژه‌های کمتری در این محدوده انرژی صرف کرده‌اند. در باقی سطوح صرف انرژی، تماماً تعداد پروژه‌ها در حدود ۵۵۰-۶۲۰ بوده.

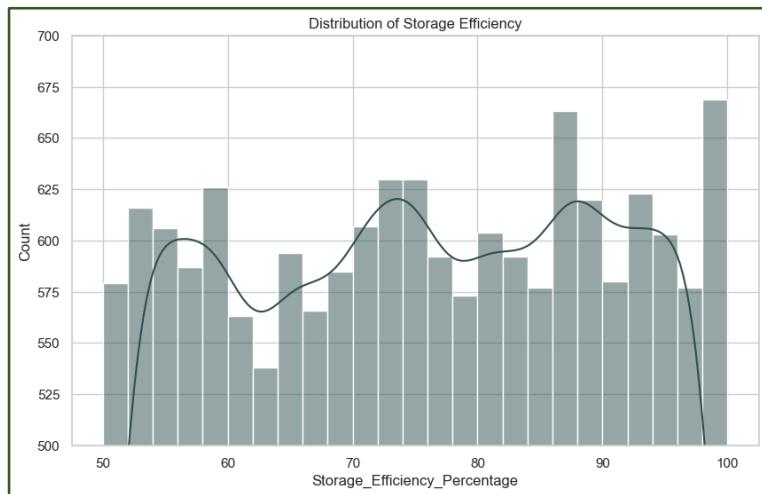


Energy Storage Capacity (MWh) (۵)

- میانگین: ۵۰۳۰,۲۰
- میانه: ۵۰۵۴,۰۴
- انحراف معیار: ۲۸۹۴,۱۰
- بازه: ۹۹۹۶,۹۴

ظرفیت ذخیره‌سازی انرژی به طور میانگین ۵۰۳۰,۲۰ مگاوات در ساعت است و میانه نیز نزدیک به میانگین است. این نشان می‌دهد که توزیع داده‌ها نسبتاً متقارن است، اما بازه گستردگی و انحراف معیار بالا نشان‌دهنده تفاوت‌های قابل توجه در ظرفیت ذخیره‌سازی پروژه‌هاست.

پراکندگی ظرفیت ذخیره سازی انرژی در پروژه‌ها در محدوده ۱۰۰۰۰--۰ مگاوات در ساعت می‌باشد و مینیمم نمودار در حدود ۱۲۰۰ الی ۱۶۰۰ اتفاق افتاده . نتیجتاً به نسبت، ظرفیت ذخیره سازی تعداد بسیار کمتری از پروژه ها در این حدود می‌باشد.



Storage Efficiency Percentage (۶)

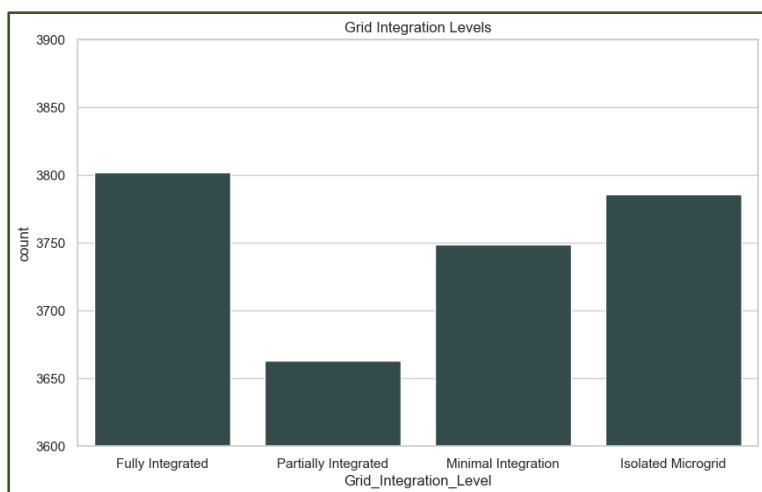
- میانگین: ۷۵,۲۲
- میانه: ۷۵,۲۸
- انحراف معیار: ۱۴,۴۹
- بازه: ۴۹,۹۹

بازدهی ذخیره‌سازی به طور میانگین ۷۵,۲۲ درصد است و میانه نزدیک به میانگین است. با این حال، بازه نشان می‌دهد که بازدهی بین ۵۰ تا ۹۹,۹۹ درصد متغیر است.

با توجه به نزدیک بودن میانه و میانگین به هم انحراف معیار پایین میتوان نتیجه گرفت بازدهی سیستم های ذخیره سازی در پروژه های مورد بررسی از پایداری نسبتاً مطلوبی برخوردار است. نقطه مینیمم نمودار در حدود ۶۲-۶۶ اتفاق افتاده که به این معناست که این مقدار بازدهی انرژی به نسبت در پروژه های کمتری اتفاق افتاده، همچنین مаксیمم نمودار در حدود ۸۶-۸۸ درصد میباشد و درصد بازدهی سیستم های ذخیره سازی انرژی پروژه ها اکثراً در این محدوده بوده است.

Grid Integration Level (۷)

میزان یکپارچگی و عدم یکپارچگی در پروژه های مورد بررسی تقریباً به یک اندازه می باشد و این باعث می شود نتوان پاسخ



دقیق و مشخصی درباره یکپارچگی پروژه ها داد و اما مطلب مورد توجه این است که تعداد کمی از پروژه ها در سطح نیمه ادغام شده قرار گرفته اند. بطور کل شهودا میتوان گفت بیشتر پروژه ها یکپارچگی ندارند و اگر دارند بسیار کم می باشد و در غیر این صورت کاملاً یکپارچه هستند.

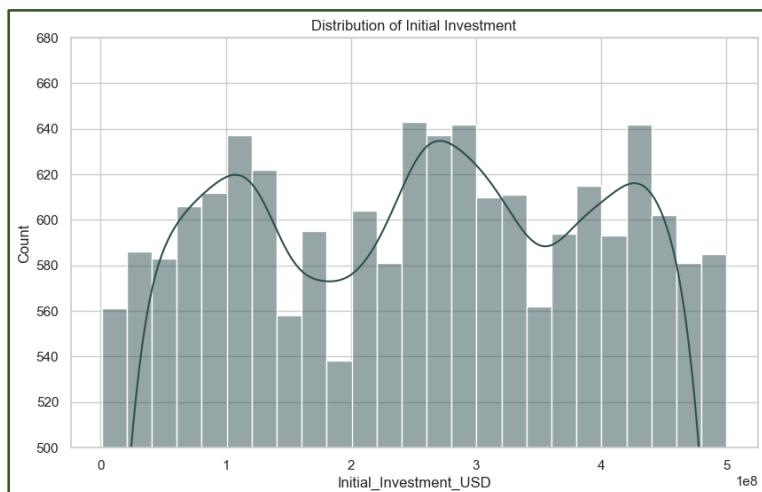
Initial Investment USD (۸)

- میانگین: ۲۵۱۴۸۴۹۱۴,۸۱

- میانه: ۲۵۳۹۹۰۹۵۴,۴۵

- انحراف معیار: ۱۴۳۲۸۴۲۷۲,۰۵

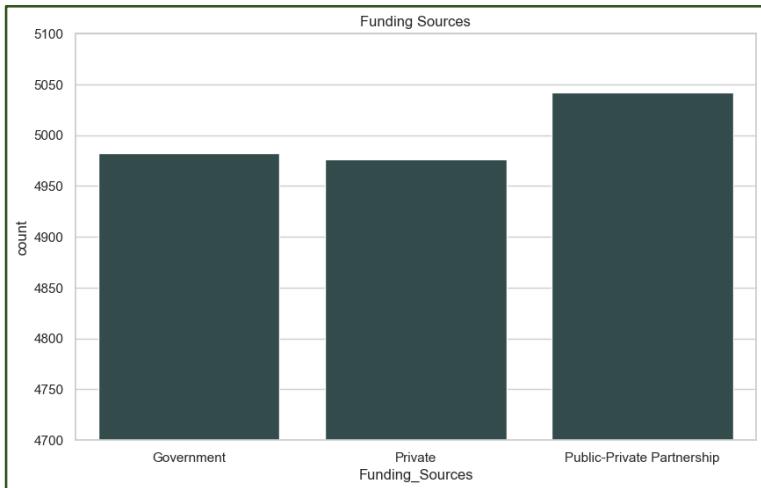
- بازه: ۴۹۸۹۳۲۶۰۴,۵۰



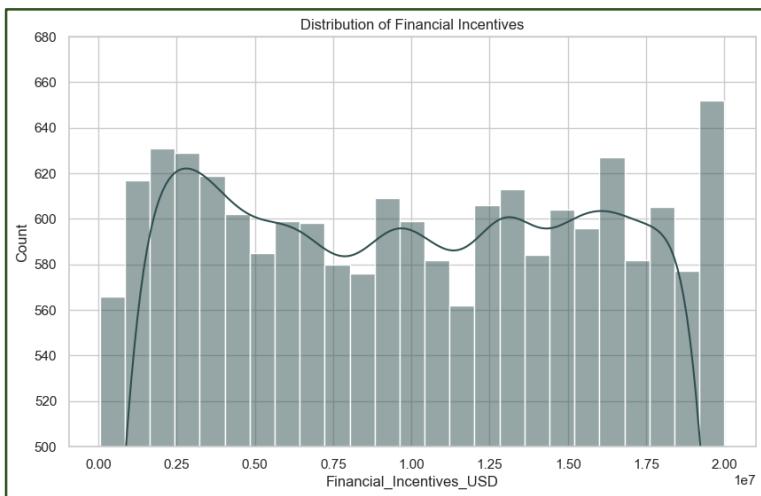
سرمایه‌گذاری اولیه به طور میانگین ۴,۸۱ ۲۵۱۴۸۴۹۱ دلار است و میانه نزدیک به میانگین است. انحراف معیار و بازه بالا نشان‌دهنده تنوع زیاد در میزان سرمایه‌گذاری اولیه در پروژه‌هاست.

میزان سرمایه‌گزاری‌های اولیه بسیار متنوع است حال آنکه بیشتر در سه حالت رخ داده و سه نقطه اپتیمم دارد که عبارتند از: ۱، ۳، ۴،۵، ۸۶۱، که بدین معناست که بیشتر سرمایه‌گزاران این مقدار سرمایه اولیه را روی پروژه‌ها گذاشته‌اند.

Funding Sources (۹)



منابع مالی برای سرمایه‌گذاران در این پروژه‌ها بیشتر به طور ترکیبی از منابع مالی دولتی و خصوصی بوده اما منابع مالی دولتی و خصوصی بطور منفرد نیز تقریباً بطور مساوی و توسط بسیاری از سرمایه‌گذاران استفاده شده‌اند.

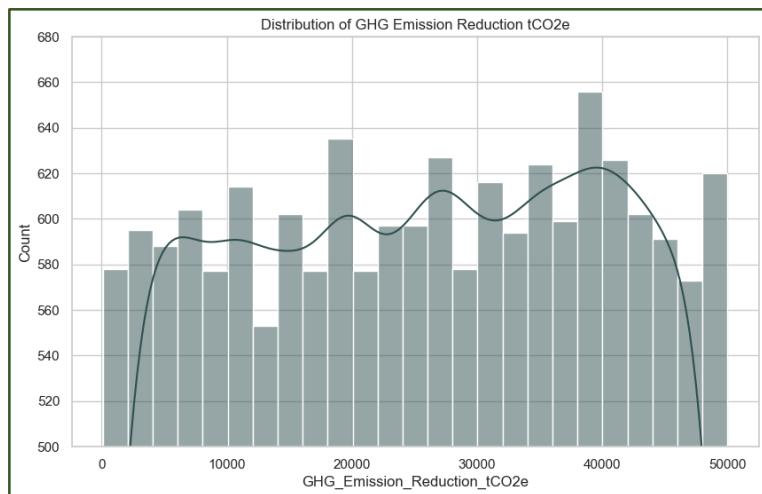


Financial Incentives (USD) (۱۳)

- میانگین: ۱۰۰۲۹۷۶۶,۲۶
- میانه: ۱۰۰۲۴۱۴۳,۵۴
- انحراف معیار: ۵۷۸۷۳۰۲,۵۸
- بازه: ۱۹۹۴۶۹۹۴,۶۱

مشوق‌های مالی به طور میانگین ۱۰۰۲۹۷۶۶,۲۶ دلار است و میانه نیز تقریباً با میانگین برابر است، اما انحراف معیار و بازه بالا نشان‌دهنده تفاوت‌های قابل توجه در میزان مشوق‌های مالی است.

مشوق های مالی با پراکندگی بالا در سطوح ۰-۲۰ میلیون دلار میباشد و با توجه به انحراف معیار و واریانس بالا نمیتوان توصیف دقیقی راجع به مشوق های مالی داشت اما نکته قابل توجه این است که مشوق های مالی در سطوح پایین تقریبا بیشتر از سطوح بالاتر میباشد و با رشد هزینه های مشوق تعداد پروژه های ذینفع کاهش پیدا میکنند.



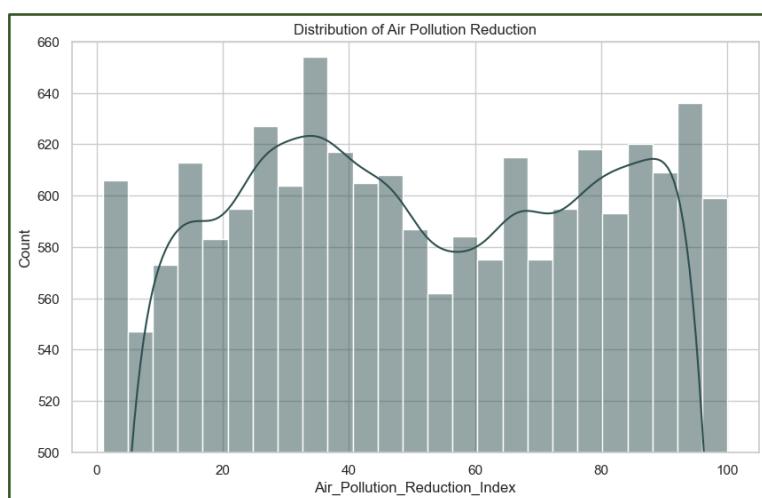
GHG Emission Reduction tCO2e (۱۱)

- میانگین: ۲۵۲۳۴,۷۲
- میانه: ۲۵۴۲۴,۴۸
- انحراف معیار: ۱۴۳۷۸,۹۲
- بازه: ۴۹۸۹۶,۶۰

کاهش انتشار گازهای گلخانه ای به طور میانگین ۲۵۲۳۴,۷۲ تن دی اکسید کربن است و میانه نزدیک به میانگین است.

بازه و انحراف معیار بالا نشان دهنده تفاوت های قابل توجه در کاهش انتشار گازهای گلخانه ای پروژه هاست.

تأثیر پروژه ها در کاهش انتشار گازهای گلخانه ای در سطوح مختلف با واریانس بالا رخ داده است و این به خودی خود پیامد بسیار خوبی برای استفاده از انرژی های تجدید پذیر میباشد. نقطه اپتیمم نمودار در حدود ۴۰۰۰۰ اتفاق افتاده است که نشان میدهد تعداد قابل توجهی از پروژه ها، در این سطح موجب کاهش انتشار گاز های گلخانه ای شده اند.

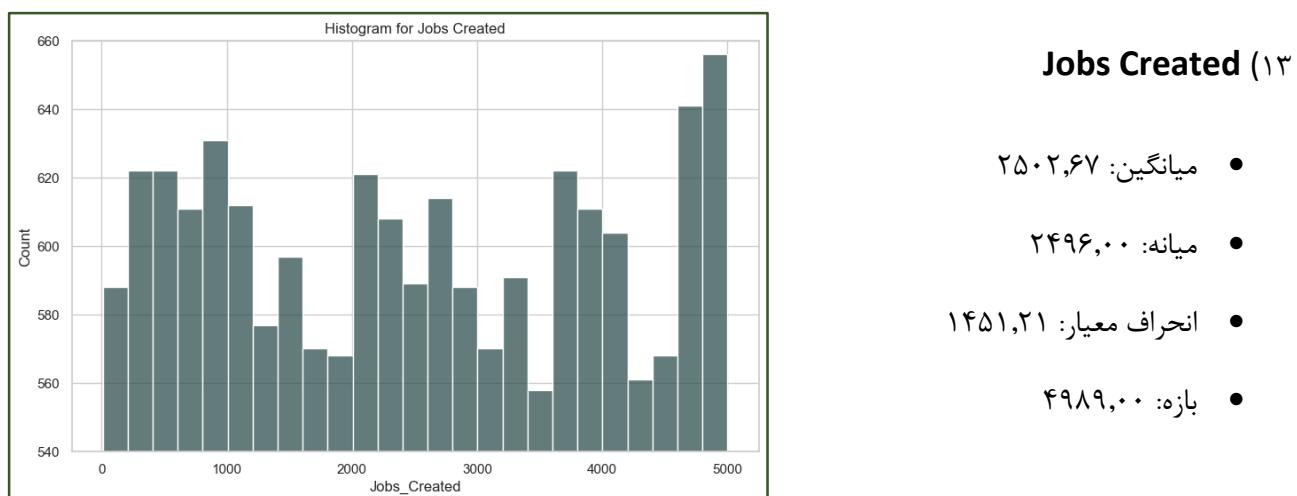


Air Pollution Reduction Index (۱۲)

- میانگین: ۵۰,۷۲
- میانه: ۵۰,۲۵
- انحراف معیار: ۲۸,۵۶
- بازه: ۹۸,۹۷

شاخص کاهش آلودگی هوا به طور میانگین ۵۰,۷۲ است و میانه نزدیک به میانگین است، اما تفاوت‌های قابل توجهی در شاخص کاهش آلودگی هوا بین پروژه‌ها وجود دارد.

پروژه‌ها در سطوح مختلف بر شاخص کاهش آلودگی هوا تاثیر گذار بوده‌اند که این موضوع به خودی خود اتفاقی مثبت تلقی می‌شود. نقاط اپتیمم نمودار در حدود ۴۰-۳۵ و ۹۰-۸۵ دیده می‌شود که نمایانگر این است که تعداد قابل توجهی از پروژه‌ها در این سطوح روی کاهش آلودگی هوا تاثیر گذار بوده‌اند و از طرفی مشاهده می‌شود که نقطه مینیمم نمودار در حدود ۶۰ می‌باشد و نسبتاً تعداد کمی از پروژه‌ها در این سطح روی شاخص کاهش آلودگی هوا موثر بوده‌اند.



تعداد شغل‌های ایجاد شده به طور میانگین ۲۵۰,۶۷ شغل است. انحراف معیار و بازه بالا نشان‌دهنده تفاوت‌های زیادی در تعداد شغل‌های ایجاد شده بین پروژه‌هاست.

مقادیر شغل‌های تولید شده در پروژه‌های مختلف متفاوت می‌باشد حال آنکه تعداد قابل توجهی از پروژه‌ها در حدود ۵۰۰۰ شغل ایجاد کرده‌اند و همچنین می‌توان مشاهده کرد گروه‌های ۱۰۰ تا ۶۱۰ از پروژه به‌طور پیوسته حدود ۱۰۰۰ شغل ایجاد کرده‌اند و این آمار و ارقام بسیار مطلوبی در حوزه کارآفرینی بوده و تاثیر بسزایی در آینده و اقتصاد خواهد داشت.

۳. پیش‌پردازش داده‌ها

۱-۳ پاکسازی داده‌ها

شناسایی و حذف ردیف‌های تکراری: ردیف‌های تکراری در دیتاست می‌توانند تحلیل را تحت تأثیر قرار دهند و منجر به نتایج نادرست شوند. بنابراین، حذف ردیف‌های تکراری ضروری است. در این مورد، هیچ ردیف تکراری در دیتاست پیدا نشد و پیغام روبه‌رو چاپ شد.

No duplicate rows found.

شناسایی داده‌های دورافتاده (Outliers): داده‌های دورافتاده می‌توانند تأثیر قابل توجهی بر تحلیل‌ها و مدل‌های آماری داشته باشند. شناسایی و در صورت لزوم حذف یا مدیریت این داده‌ها می‌تواند دقت تحلیل‌ها را افزایش دهد. برای شناسایی داده‌های دورافتاده از روش Interquartile Range (IQR) استفاده شده است. در این روش:

چارک اول (Q1) و چارک سوم (Q3) برای هر ستون محاسبه می‌شود. سپس محدوده بین چارکی (IQR) محاسبه می‌شود. حد پایین و حد بالا به ترتیب به صورت $\text{lower_bound} = Q1 - 1.5 * \text{IQR}$ و $\text{upper_bound} = Q3 + 1.5 * \text{IQR}$ محاسبه می‌شوند. داده‌هایی که خارج از این حدود باشند به عنوان دورافتاده شناسایی می‌شوند.

طبق خروجی که در زیر مشاهده می‌شود برای هیچ یک از ستون‌های عددی، داده‌های دورافتاده شناسایی نشدنند. همچنین می‌توان با کشیدن نمودار جعبه‌ای نشان داد که هیچ کدام از متغیرها داده دور افتاده ندارند.

```
No outliers found in Installed_Capacity_MW.  
No outliers found in Energy_Production_MWh.  
No outliers found in Energy_Consumption_MWh.  
No outliers found in Energy_Storage_Capacity_MWh.  
No outliers found in Storage_Efficiency_Percentage.  
No outliers found in Initial_Investment_USD.  
No outliers found in Financial_Incentives_USD.  
No outliers found in GHG_Emission_Reduction_tCO2e.  
No outliers found in Air_Pollution_Reduction_Index.  
No outliers found in Jobs_Created.
```

عدم وجود ردیف‌های تکراری و عدم وجود داده‌های دورافتاده نشان‌دهنده این است که داده‌ها از کیفیت مناسبی برخوردار هستند و می‌توان با اطمینان بیشتری به تحلیل و مدل‌سازی روی آن‌ها پرداخت.

۲-۳ داده‌های گمشده

در این **dataframe** هیچ داده‌ی گمشده‌ای وجود ندارد، زیرا تعداد داده‌های غیر گمشده در هر ستون برابر با تعداد کل ردیف‌ها (۱۵۰۰۰) است. یعنی ساختار دیتافریم به خوبی تعریف شده و نیازی به اقدامات اصلاحی برای داده‌های گمشده نیست.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15000 entries, 0 to 14999
Data columns (total 13 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   Type_of_Renewable_Energy    15000 non-null   category
 1   Installed_Capacity_MW     15000 non-null   float64
 2   Energy_Production_MWh    15000 non-null   float64
 3   Energy_Consumption_MWh   15000 non-null   float64
 4   Energy_Storage_Capacity_MWh 15000 non-null   float64
 5   Storage_Efficiency_Percentage 15000 non-null   float64
 6   Grid_Integration_Level    15000 non-null   category
 7   Initial_Investment_USD    15000 non-null   float64
 8   Funding_Sources          15000 non-null   category
 9   Financial_Incentives_USD 15000 non-null   float64
 10  GHG_Emission_Reduction_tCO2e 15000 non-null   float64
 11  Air_Pollution_Reduction_Index 15000 non-null   float64
 12  Jobs_Created            15000 non-null   int64  
dtypes: category(3), float64(9), int64(1)
memory usage: 1.2 MB
```

۳-۳ استانداردسازی داده‌ها

استانداردسازی (Standardization) یکی از مراحل پاکسازی داده‌های است که به دلیل زیر انجام آن روی داده‌های انرژی، حیاتی و مهم است.

مقایسه‌پذیری: مقادیر متغیرها در دیتاست ممکن است در مقیاس‌های مختلفی باشند. برای مثال، برخی ویژگی‌ها ممکن است در مقیاس صدها یا هزارها باشند مثل (Installed Capacity MW)، در حالی که دیگر ویژگی‌ها ممکن است در

مقیاس دهها یا صدها باشند مثل (Storage Efficiency Percentage) استانداردسازی به ما اجازه می‌دهد تا همه ویژگی‌ها را در یک مقیاس مشترک قرار دهیم و مقایسه‌پذیری آن‌ها را افزایش دهیم.

برای استانداردسازی داده‌ها از کلاس StandardScaler در کتابخانه scikit-learn استفاده شده است. در نتیجه نهایی، تمامی مقادیر ستون‌های عددی به یک مقیاس استاندارد (میانگین صفر و انحراف معیار یک) تبدیل شدن، که باعث می‌شود داده‌ها برای مدل‌سازی آماده‌تر باشند و از مشکلات مقیاسی جلوگیری شود.

۴-۳ تبدیل متغیرها

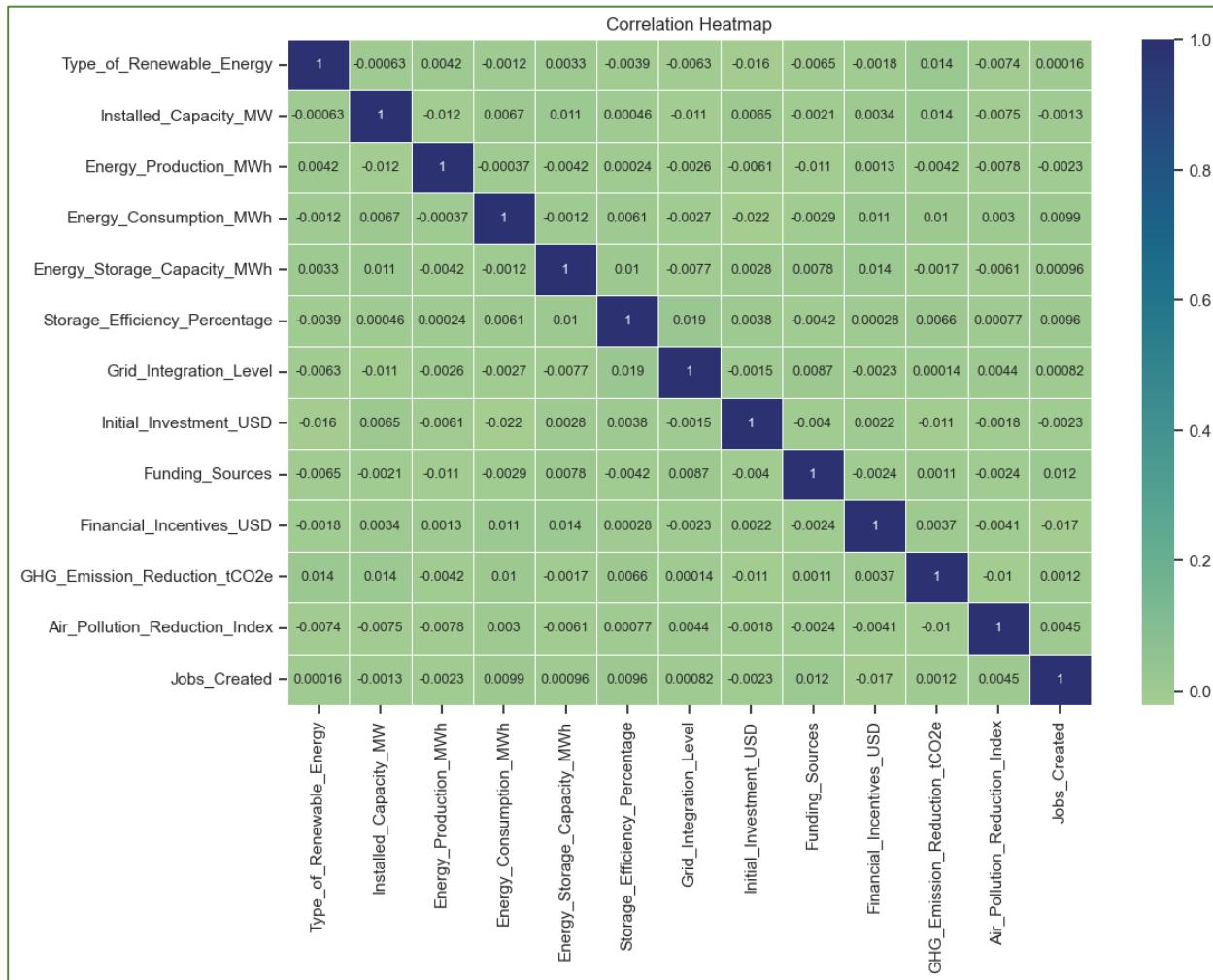
در این پژوهش تبدیلات برای نرمال کردن توزیع متغیرها امکان پذیر نبود. این مشکل میتواند دلایل مختلفی داشته باشد از جمله:

- **دامنه محدود داده‌ها:** اگر دامنه داده‌ها به طور تنوعی محدود شده باشد یعنی اعداد فقط در یک بازه خاص با اختلاف کم قرار گرفته باشند، توزیع داده‌ها ممکن است نرمال نباشد.
 - **توزیع‌های ذاتاً غیر نرمال:** برخی توزیع‌ها ذاتاً به گونه‌ای هستند که نمی‌توانند به توزیع نرمال تبدیل شوند. به عنوان مثال، توزیع‌های گسسته مانند توزیع پواسون یا توزیع‌های بسیار چولگی دار مثل پواسون، تی، توانی و کوشی ممکن است با هیچ تبدیلی به توزیع نرمال تبدیل نشوند.
- با توجه به محدودیت مشاهده شده از نرمال کردن توزیع‌ها صرف نظر شد.

۴. الگوهای پنهان و همبستگی‌ها

۱-۴ ماتریس همبستگی

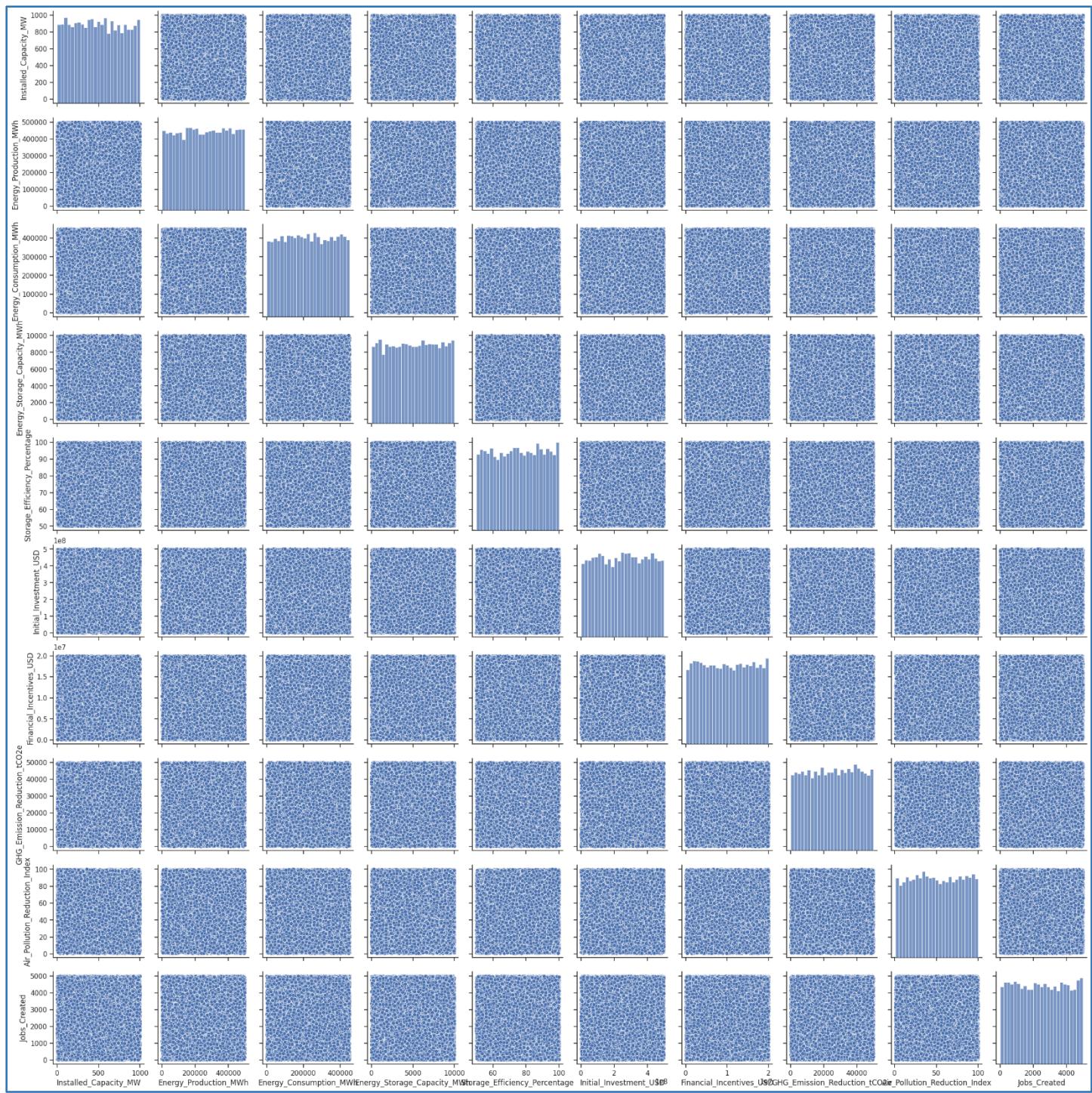
بعد از ساختن ماتریس همبستگی نتایج آن را با کمک نمودار حرارتی زیر بررسی می‌کنیم.



جدول ارائه شده، ماتریس همبستگی بین ۱۳ متغیر مربوط به منابع انرژی تجدیدپذیر را نشان می‌دهد. اعداد داخل سلول‌ها ضریب همبستگی پیرسون را نمایش می‌دهند که مقدار آن‌ها بین -۱ تا ۱ است. رنگ آبی تیره نشان‌دهنده همبستگی بالا و سبز روشن، نشان‌دهنده همبستگی پایین هستند. مقادیر نزدیک به ۱ یا -۱ نشان‌دهنده همبستگی قوی هستند و مقادیر نزدیک به ۰ نشان‌دهنده همبستگی ضعیف یا عدم همبستگی هستند.

بیشترین همبستگی بین سطح یکپارچگی و کارایی سیستم‌های ذخیره‌سازی است که مقدار ۱۹٪ را دارد و تا حدودی به صفر نزدیک است. با توجه به نتایج می‌توان گفت متغیرها دو به دو از همدیگر مستقل هستند و همبستگی خیلی کمی دارند. برای بررسی دقیق‌تر باید از تست‌های آماری استفاده شود.

۲-۴ نمودار پراکندگی



نمودار پراکندگی نوعی نمودار است که برای نمایش رابطه بین دو متغیر به کار می‌رود. در این نمودار، مقادیر یک متغیر در محور افقی و مقادیر متغیر دیگر در محور عمودی رسم می‌شوند. هر نقطه در نمودار نشان دهنده یک جفت از مقادیر مربوط به دو متغیر است.

همبستگی مثبت: اگر نقاط در نمودار پراکندگی از سمت چپ پایین به سمت راست بالا شیب داشته باشند، نشان دهنده همبستگی مثبت بین دو متغیر است. این بدان معناست که با افزایش مقادیر یک متغیر، مقادیر متغیر دیگر نیز به طور کلی افزایش می‌یابد.

همبستگی منفی: اگر نقاط در نمودار پراکندگی از سمت چپ بالا به سمت راست پایین شیب داشته باشند، نشان دهنده همبستگی منفی بین دو متغیر است. این بدان معناست که با افزایش مقادیر یک متغیر، مقادیر متغیر دیگر به طور کلی کاهش می‌یابد.

عدم همبستگی: اگر نقاط در نمودار پراکندگی به طور تصادفی در سراسر نمودار پخش شده باشند، نشان دهنده عدم وجود همبستگی بین دو متغیر است. این بدان معناست که هیچ رابطه قابل توجهی بین مقادیر دو متغیر وجود ندارد.

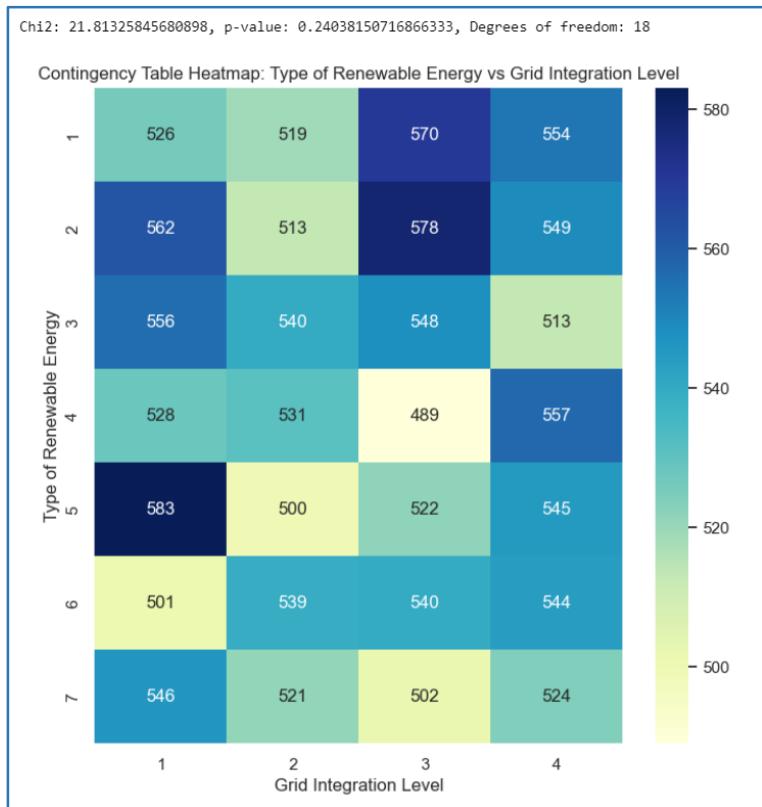
کشیدن این نمودارها به ما عدم همبستگی متغیرها را نشان می‌دهد و همان نتیجه‌های که از ماتریس همبستگی گرفتیم در این قسمت تکرار می‌شود.

با توجه به دو نمودار بالا، نزدیک صفر بودن اغلب ضرایب همبستگی و همچنین عدم مشاهده هیچ رابطه خطی، می‌توان نتیجه گرفت رابطه غیرخطی بین متغیرها برقرار است که در این تحقیق، تحلیل و بررسی این مسئله مد نظر نبوده و خارج از محدوده این پژوهش است.

۳-۴ آزمون‌های آماری

آزمون خی-دو (Chi-Square) یک آزمون آماری است که برای بررسی ارتباط بین دو متغیر رسته‌ای استفاده می‌شود. این آزمون به ما کمک می‌کند تا بفهمیم آیا بین دو متغیر رابطه‌ای وجود دارد یا خیر. با استفاده از جدول پیشاپندی (Contingency Table) که فراوانی‌های مشترک دو متغیر را به ما نشان می‌دهد، این آزمون را انجام می‌دهیم.

• ارتباط بین نوع انرژی تجدیدپذیر و سطح یکپارچه‌سازی شبکه



مقدار آماره Chi2 برابر با ۲۱,۸۱ است. مقدار p-value

برابر با ۰,۲۴۰۳ است. درجه آزادی برابر با ۱۸ است.

مقدار p-value بزرگتر از ۰,۰۵ است پس فرض H0 رد نمی‌شود. فرض صفر در اینجا بیان می‌کند که هیچ رابطه‌ای بین نوع انرژی تجدیدپذیر و سطح یکپارچگی شبکه وجود ندارد.

به عبارت دیگر، این نتیجه نشان می‌دهد که هیچ شواهد آماری قوی‌ای برای ارتباط بین این دو متغیر وجود ندارد و می‌توان گفت یکپارچگی شبکه به نوع انرژی بستگی ندارد.

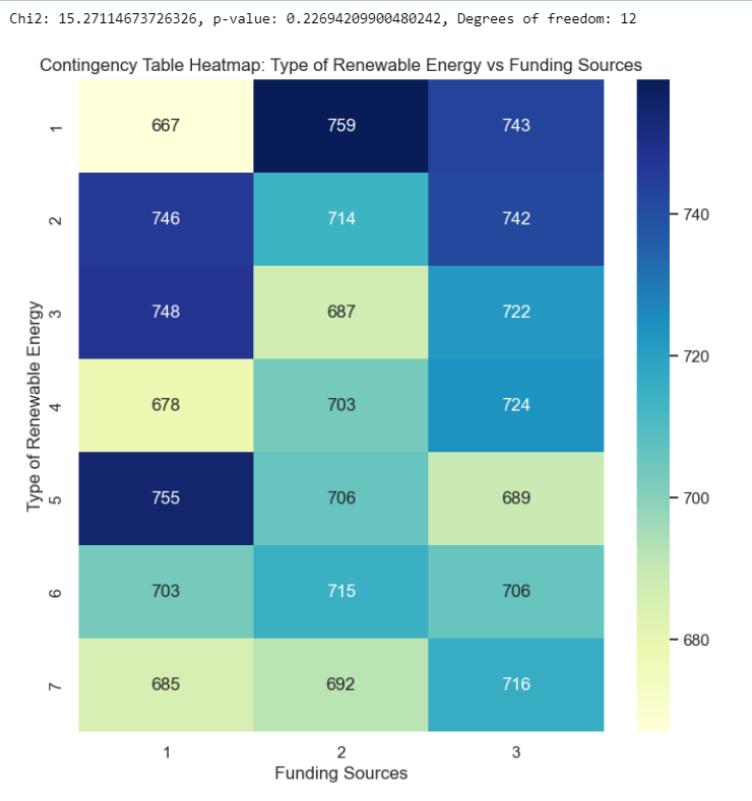
نمودار حرارتی به صورت بصری نشان می‌دهد که داده‌ها در ترکیب‌های مختلف نوع انرژی تجدیدپذیر و سطح یکپارچگی شبکه چگونه توزیع شده‌اند. این نمودار می‌تواند به شناسایی الگوهای خاصی در داده‌ها کمک کند، اما نتایج تست آماری Chi-Square همچنان بیان می‌کند که این الگوها احتمالاً به صورت تصادفی ایجاد شده‌اند و ارتباط معنی‌داری بین این دو متغیر وجود ندارد.

• ارتباط بین نوع انرژی تجدیدپذیر و منابع مالی

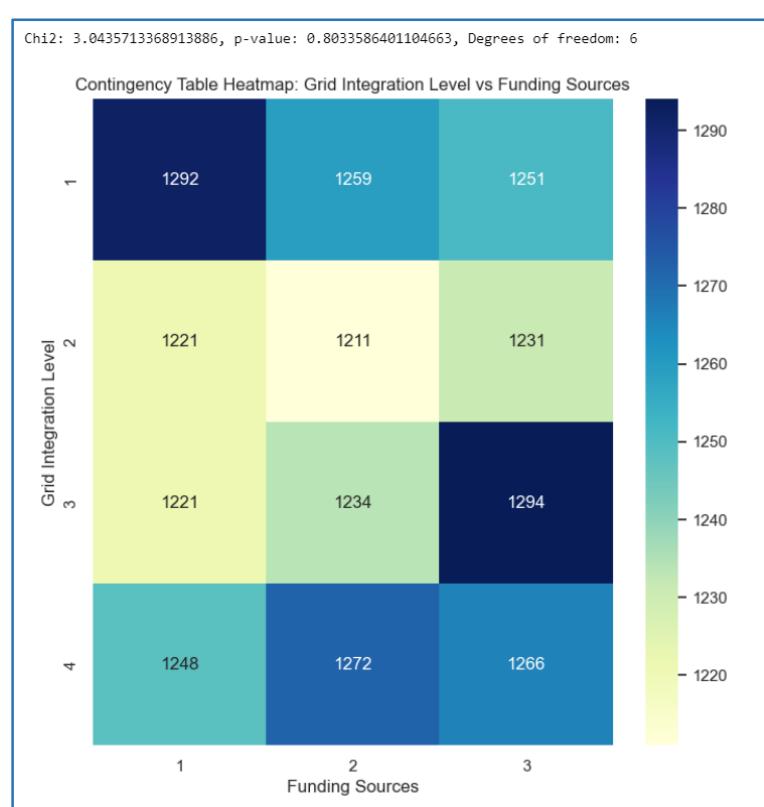
مقدار آماره Chi2 برابر با ۱۵,۲۷ است.

مقدار p-value برابر با ۰,۲۲۶ است.

درجه آزادی برابر با ۱۲ است.



مقدار $p\text{-value}$ بزرگتر از 0.05 است یعنی نمی‌توان فرض صفر را رد کرد. فرض صفر در اینجا بیان می‌کند که هیچ رابطه‌ای بین نوع انرژی تجدیدپذیر و منابع تامین مالی وجود ندارد. به عبارت دیگر، این نتیجه نشان می‌دهد که هیچ شواهد آماری قوی‌ای برای ارتباط بین این دو متغیر وجود ندارد.



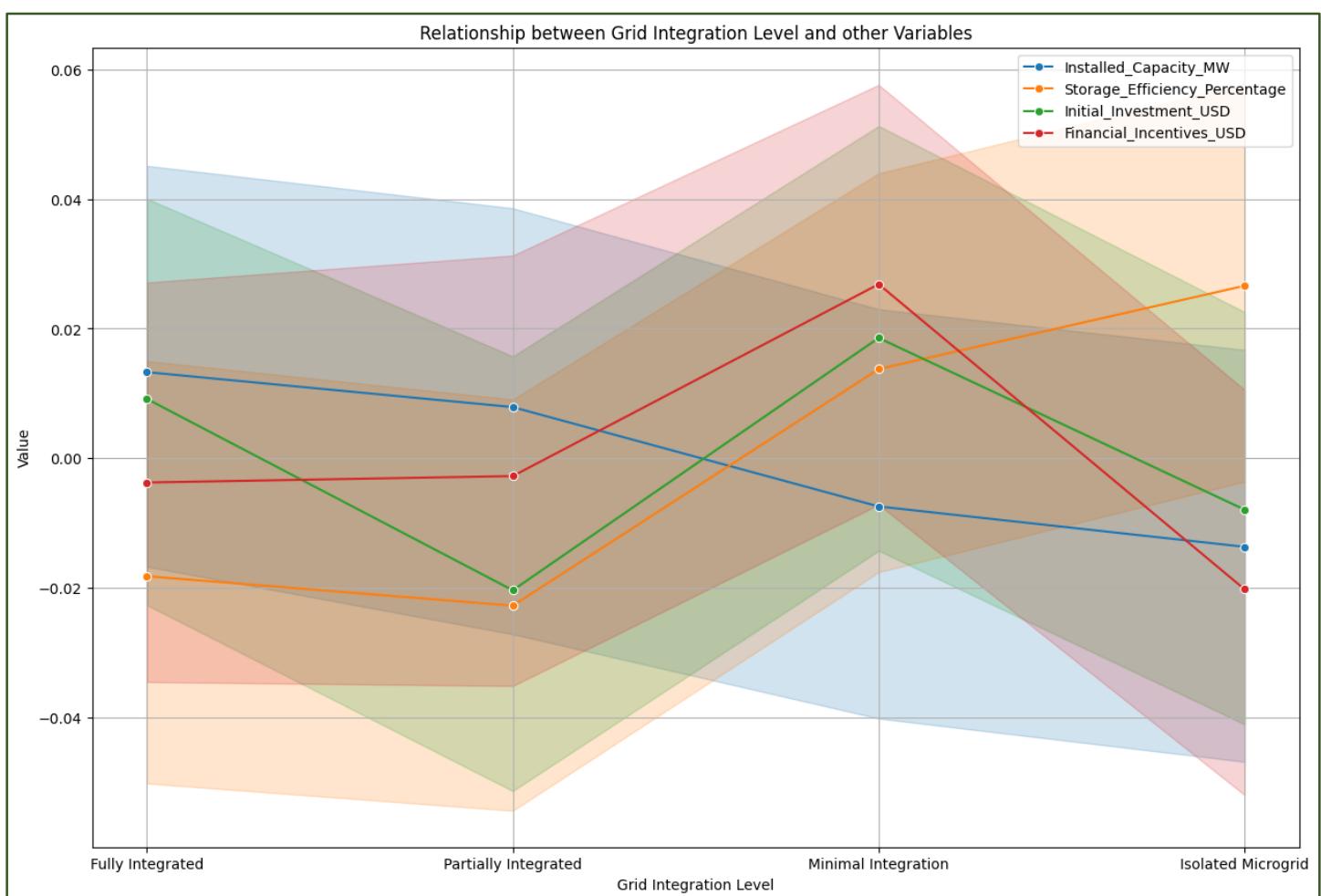
مقدار Chi2 برابر با ۳.۰۴ است. مقدار $p\text{-value}$ برابر با 0.803 است. درجه آزادی برابر با ۶ است. مقدار $p\text{-value}$ بسیار بزرگتر از 0.05 است یعنی H_0 رد نمی‌شود. فرض صفر در اینجا بیان می‌کند که هیچ رابطه‌ای بین سطح یکپارچه‌سازی شبکه و منابع تامین مالی وجود ندارد. به عبارت دیگر، این نتیجه نشان می‌دهد یکپارچه بودن شبکه و منبع مالی از هم مستقل هستند.

۵. بصری سازی نتایج

۱-۵ نمودار خطی

برای رسم این نمودارها مقادیر عددی داده انرژی استاندارد شده‌اند تا بتوانیم در یک مقیاس معین به بررسی روابط آن‌ها پردازیم. میانگین متغیر مورد نظر در سطوح مختلف محاسبه می‌شود و با دایره نشان داده شده، سایه‌های روی نمودار نیز نشان‌دهنده بازه اطمینان و تغییرات آن متغیر است.

• رابطه بین سطح یکپارچگی و بقیه متغیرها



سطح یکپارچگی پروژه‌ها و ظرفیت نصب شده پروژه‌ها :

با توجه به نمودار مشاهده می‌شود که هرچه یکپارچگی کاهش پیدا می‌کند ظرفیت نصب پروژه‌ها نیز اکیدا کاهش پیدا می‌کند. این به این معناست که عموماً تمایل به سمت نصب پروژه‌های یکپارچه می‌باشد.

سطح یکپارچگی و میزان سرمایه گذاری اولیه:

با توجه به نمودار می‌توان متوجه دو نقطه اپتیمم شد که در سطوح پروژه‌های کاملاً یکپارچه و پروژه‌هایی با یکپارچگی جزئی اتفاق افتاده است. این بدان معناست که عموماً سرمایه گذاران روی چنین پروژه‌هایی سرمایه گذاری می‌کنند و در صد پایینی از سرمایه گذاری‌های اولیه روی پروژه‌هایی که تقریباً یکپارچه هستند انجام می‌شود.

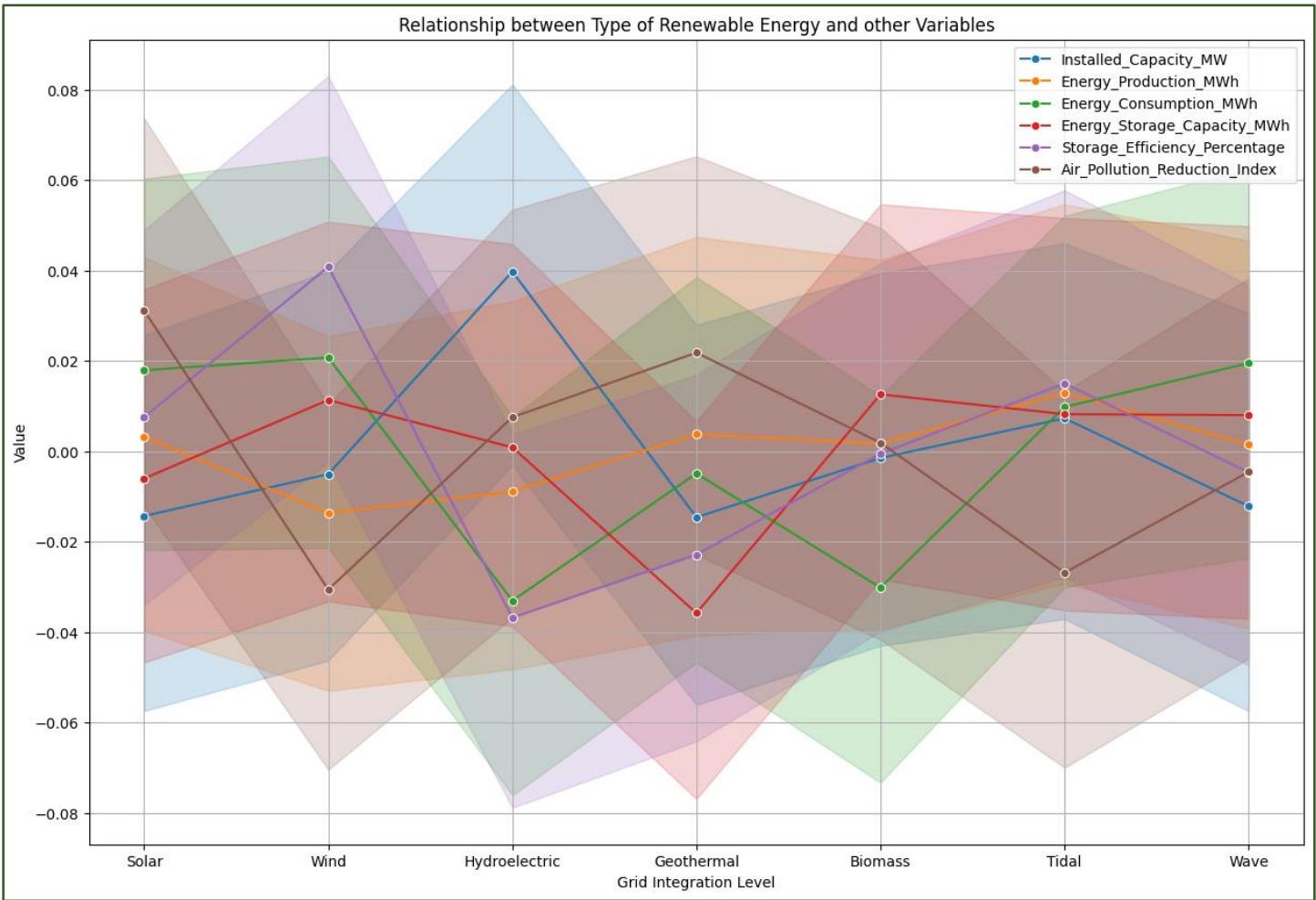
در صد بازدهی سیستم و یکپارچگی:

بازدهی سیستم‌ها با کاهش یکپارچگی بطور قابل توجهی افزایش پیدا می‌کند. در پروژه‌های تقریباً یکپارچه بازدهی سیستم‌های ذخیره‌سازی مینیمم مقدار را دارد و با کاهش یکپارچگی بازدهی سیستم‌های ذخیره‌سازی به طور قابل توجهی افزایش پیدا می‌کند. بطور کل می‌توان گفت در پروژه‌هایی با عدم یکپارچگی و بخصوصی پروژه‌هایی با یکپارچگی جزئی، بازدهی سیستم‌های ذخیره‌سازی در بهترین حالت قرار دارند.

مشوق‌های مالی و یکپارچگی:

نقشه ماکسیمم نمودار میزان سرمایه گذاری مشوق‌های مالی و یکپارچگی در سطح پروژه‌هایی با یکپارچگی جزئی اتفاق افتاده که نشان می‌دهد عموماً مشوق‌های مالی روی چنین پروژه‌هایی سرمایه گذاری کرده‌اند و مینیمم سرمایه گذاری مشوق‌های مالی در سطح پروژه‌هایی با عدم یکپارچگی اتفاق افتاده و چنین پروژه‌هایی مطلوب مشوقین مالی نمی‌باشد.

• رابطه بین انواع انرژی و بقیه متغیرها



ظرفیت پروژه‌های نصب شده و انواع انرژی‌های تجدیدپذیر:

با توجه به نمودار می‌توان دریافت بیشتر پروژه‌های نصب شده روی انرژی تجدیدپذیر هیدرولکتریک بوده و مینیمم پروژه‌ها روی انرژی زیست‌توده نصب شده‌اند، پروژه‌های نصب شده روی مابقی انرژی‌های تجدیدپذیر تقریباً به اندازه‌ی یکسانی است.

میزان تولید انرژی و انواع انرژی‌های تجدیدپذیر:

سطح تولیدات انرژی در انواع انرژی‌های تجدیدپذیر تفاوت چندانی با هم ندارند، با این حال انرژی تجدیدپذیر جزر و مد به نسبت مابقی انرژی‌ها سطح تولیدات انرژی بیشتری داشته و مینیمم تولیدات انرژی مربوط به انرژی باد میباشد. انرژی‌های زیست توده، زمین‌گرمایی و موج سطح تولیدات انرژی یکسانی دارند.

ظرفیت ذخیره سازی انرژی و انواع انرژی‌های تجدیدپذیر:

انرژی زیست توده بطور قابل توجهی میزان کمتری انرژی نسبت به مابقی انرژی‌های تجدیدپذیر ذخیره می‌کند و انرژی‌های بادی و زمین‌گرمایی بیشترین ذخیره انرژی را دارند. همچنین میزان ذخیره انرژی در انواع انرژی‌های جزر و مد و موج تقریباً یکسان و مطلوب می‌باشد.

میزان مصرف انرژی و انواع انرژی‌های تجدیدپذیر :

انرژی‌های خورشیدی، باد و موجی بیشترین میزان مصرف انرژی را دارند و مینیمم نمودار در انواع انرژی هیدروالکتریک و زیست توده مشاهده می‌شود و این نوع انرژی‌های تجدیدپذیر نسبت به مابقی میزان کمترین انرژی مصرف می‌کند.

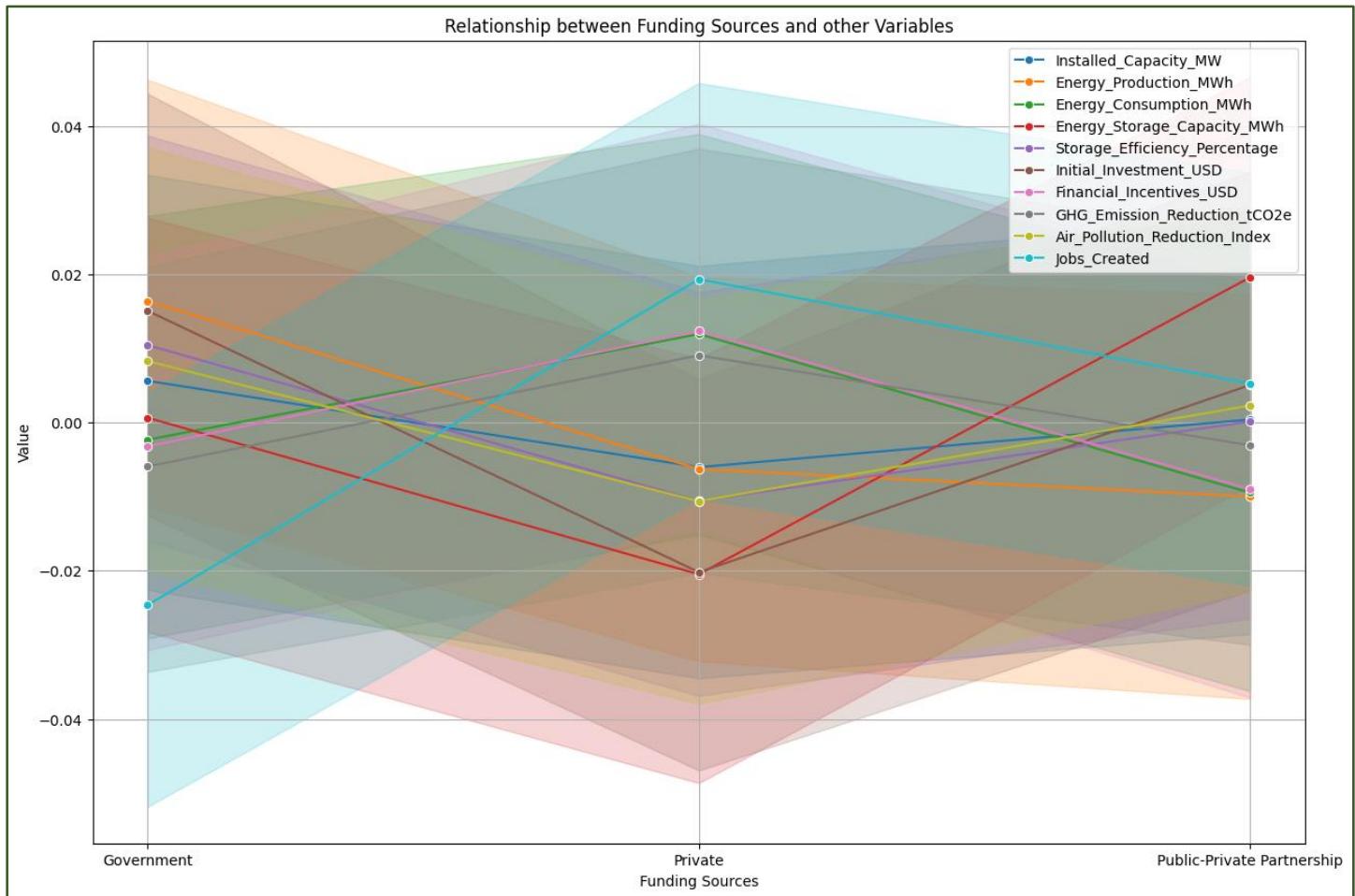
درصد بازدهی سیستم‌های ذخیره‌سازی در انواع انرژی‌های تجدیدپذیر:

بیشترین میزان بازدهی سیستم‌های ذخیره‌سازی انرژی با اختلاف زیادی در نوع انرژی باد مشاهده می‌شود و مینیمم بازدهی سیستم‌های ذخیره‌سازی انرژی مربوط به انرژی هیدروالکتریک می‌باشد. انرژی‌های خورشیدی و جزر و مدی و همچنین انرژی‌های موجی و زیست توده‌ای دو به دو بازدهی‌های یکسانی دارند.

میزان کاهش آلودگی هوا و انواع انرژی‌های تجدیدپذیر:

با توجه به نمودار انرژی بادی و جزر و مد، کمترین میزان تاثیر را بر روی کاهش آلودگی هوا دارند و انرژی خورشیدی بیشترین میزان تاثیر بر روی کاهش آلودگی هوا را دارد.

• رابطه بین منبع تامین مالی و بقیه متغیرها



ظرفیت نصب پروژه ها و انواع منابع مالی:

ظرفیت نصب پروژه ها در حالتی که دولت منابع مالی را فراهم کرده بیشترین مقدار را داشته و کمترین مقدار نصب پروژه ها مربوط به حالتی است که منابع مالی شخصی سرمایه گذاری کرده اند.

میزان تولید انرژی و منابع مالی:

با توجه به اینکه میزان نصب پروژه‌ها به هنگام دولتی بودن منابع مالی بیشترین مقدار را داشته بنظر منطقی می‌رسد که میزان تولیدات انرژی نیز در این حالت بیشترین مقدار را داشته باشد. همچنین کمترین مقدار تولیدات انرژی مرتبط با زمانی است که از منبع مالی ترکیبی (دولتی و شخصی) استفاده شده است.

میزان مصرف انرژی و منابع مالی:

میزان مصرف انرژی در حالتی که از منابع مالی شخصی استفاده شده بیشترین مقدار را دارد و همچنین در حالاتی که منابع مالی دولتی و یا ترکیبی است میزان مصرف انرژی تقریباً یکسان می‌باشد.

ظرفیت ذخیره‌سازی انرژی و منابع مالی:

میزان ذخیره‌سازی انرژی در حالتی که از منبع مالی ترکیبی استفاده شده ماقسیمم مقدار را دارد و مینیمم میزان ذخیره‌سازی انرژی مربوط به زمانیست که از منابع مالی شخصی برای اجرای پروژه‌ها استفاده می‌شود.

درصد بازدهی سیستم‌های ذخیره‌سازی و منابع مالی:

بازدهی سیستم‌های ذخیره‌سازی انرژی به هنگام بهره برداری از منابع مالی دولتی بیشترین مقدار را داشته و هنگامی که منابع مالی شخصی باشد می‌توان دید بازدهی سیستم‌های ذخیره‌سازی مینیمم می‌باشد.

میزان سرمایه‌گذاری اولیه و منابع مالی:

میزان سرمایه‌گذاری‌های اولیه به هنگام دولتی بودن منابع مالی بطور قابل توجهی بیشتر است و مینیمم مقدار را وقتی می‌گیرد که از منابع مالی شخصی استفاده شده باشد.

میزان مشوق‌های مالی و منابع :

میزان مشوق‌های مالی به هنگام استفاده از منابع مالی شخصی بیشترین مقدار است و در صورت استفاده از منابع مالی دولتی و ترکیبی میزان مشوق‌های مالی تقریباً یکسان است.

میزان کاهش گازهای گلخانه‌ای و منابع مالی:

می‌توان مشاهده کرد که در صورت استفاده از منابع مالی شخصی میزان کاهش گازهای گلخانه‌ای بیشترین مقدار را داشته و در صورت بهره‌بری از منابع مالی دولتی و یا ترکیبی، میزان کاهش گازهای گلخانه‌ای تقریباً یکسان است.

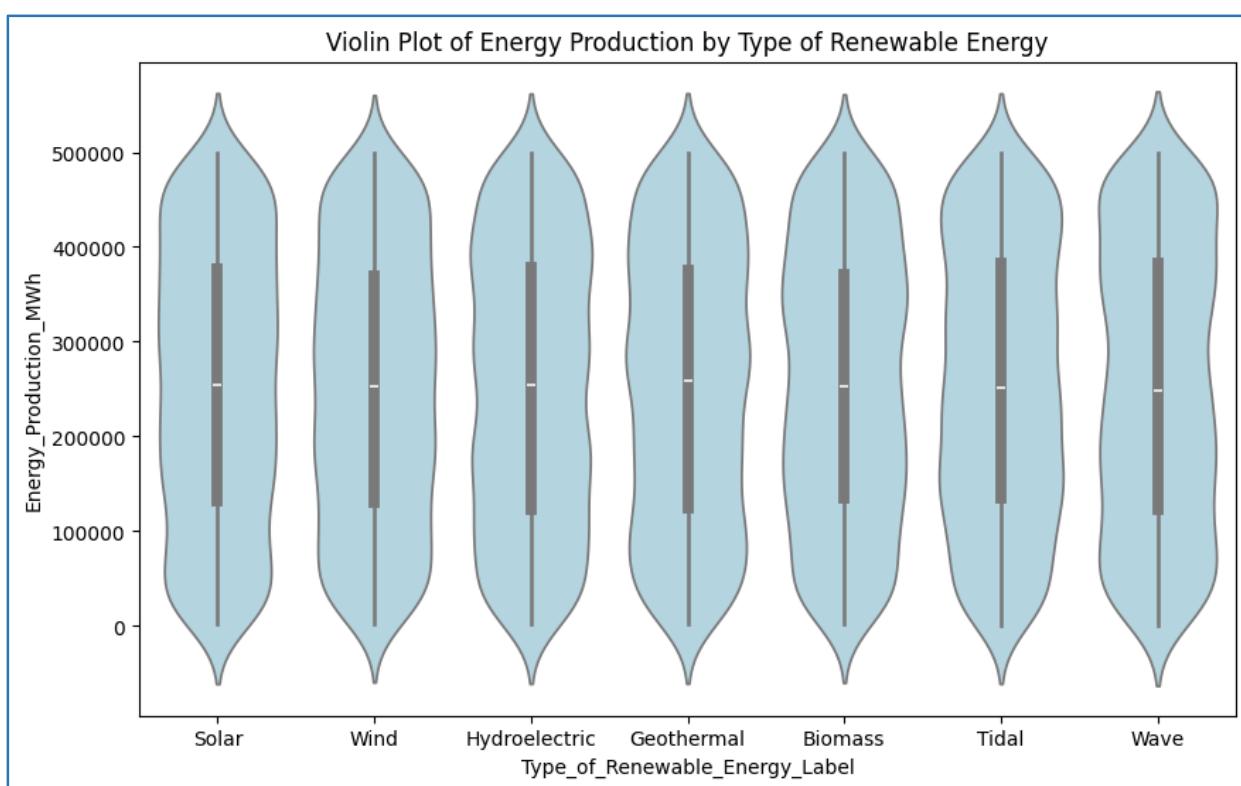
میزان کاهش آلودگی هوا و منابع مالی:

برخلاف میزان کاهش گازهای گلخانه‌ای، میزان کاهش آلودگی هوا در صورت استفاده از منابع مالی شخصی مینیمم مقدار را داشته و ماکزیمم مقدار کاهش آلودگی هوا به هنگام استفاده از منابع مالی دولتی اتفاق می‌افتد.

تعداد شغل‌های ایجاد شده و منابع مالی:

تعداد شغل‌های ایجاد شده در اجرای پروژه با منابع مالی شخصی به طور قابل توجهی بیشتر از دیگر حالات اجرای پروژه می‌باشد و مینیمم تعداد شغل‌های ایجاد شده به هنگام استفاده از منابع مالی دولتی می‌باشد.

نمودار ویولنی ۲-۵



در این بخش، نمودار ویولن (Violin Plot) تولید انرژی بر اساس نوع انرژی تجدیدپذیر را رسم شده است. این نمودار ترکیبی از نمودار جعبه‌ای (Box Plot) و نمودار چگالی (KDE: Kernel Density Estimate) است و اطلاعاتی در مورد توزیع داده‌ها به ما می‌دهد. محور افقی، نوع انرژی تجدیدپذیر و محور عمودی نمودار، مقدار تولید انرژی را در مگاوات بر ساعت نشان می‌دهد.

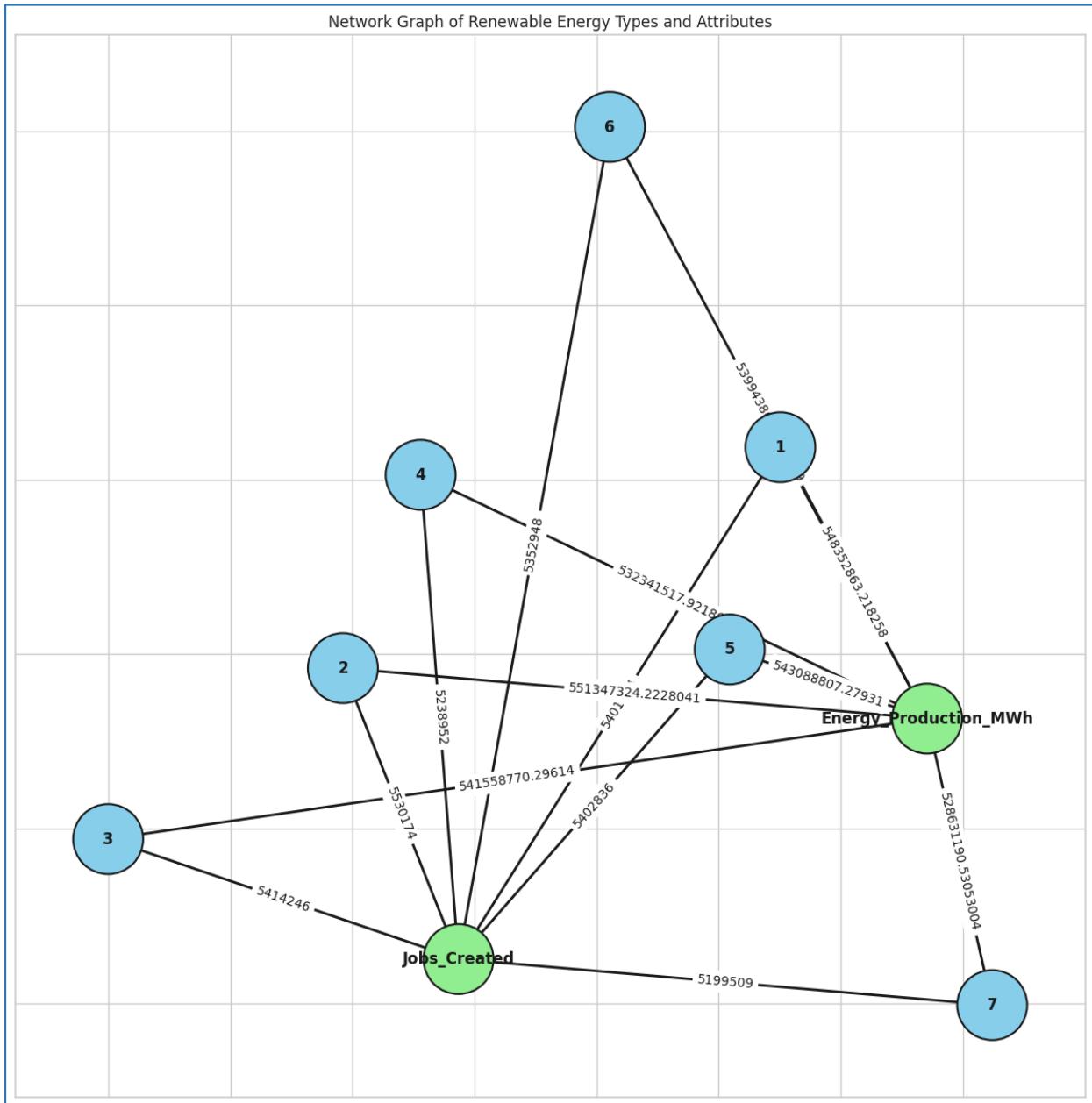
این نمودار به شکل یک ویولن (یا دو ویولن رو برو هم) که نمایانگر چگالی توزیع داده‌ها می‌باشد، است. هر چه پهنه‌ای ویولن بیشتر باشد، تعداد داده‌ها در آن محدوده بیشتر است. همچنین نمودار جعبه‌ای درون ویولن قرار دارد و نشان‌دهنده چارک‌های اول، دوم (میانه)، و سوم داده‌ها است. این شکل نشان می‌دهد که میانه‌های تولید انرژی در انواع مختلف انرژی تقریباً در یک راستا قرار دارند و تفاوت چشمگیری مشاهده نمی‌شود.

این نمودار به ما کمک می‌کند تا توزیع داده‌ها برای هر نوع انرژی تجدیدپذیر را به خوبی مشاهده کنیم و می‌توانیم توزیع تولید انرژی برای انواع مختلف انرژی تجدیدپذیر را با یکدیگر مقایسه کنیم. همچنین اگر داده‌هایی خارج از محدوده‌های معمولی توزیع باشند، می‌توانیم آن‌ها را شناسایی کنیم که در اینجا داده دورافتاده‌ای مشاهده نمی‌شود.

۳-۵ نمودار شبکه‌ای

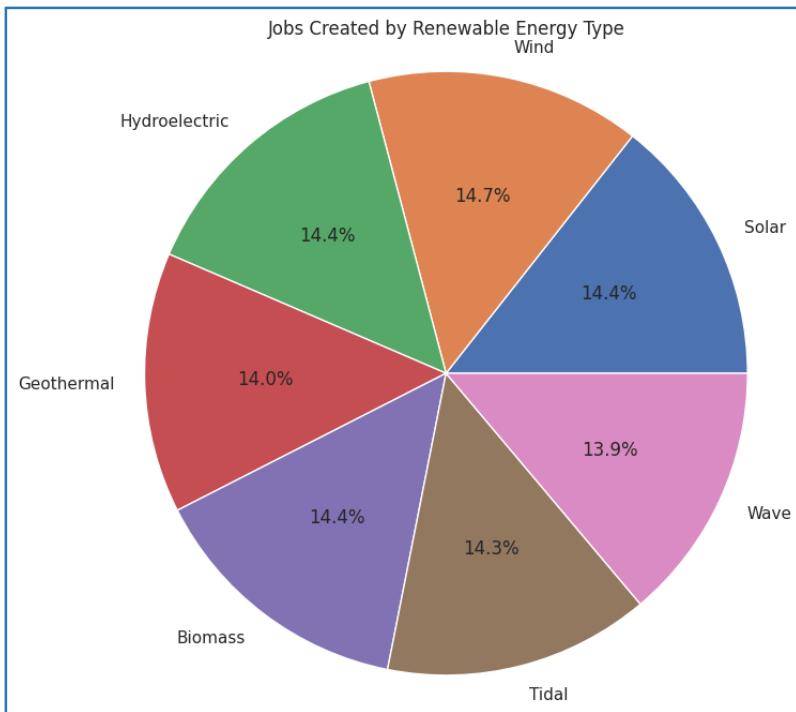
در این قسمت نمودار شبکه‌ای انواع انرژی تجدیدپذیر و دو ویژگی مهم (مشاغل ایجاد شده و تولید انرژی) رسم می‌شود. هدف از رسم این نمودار بررسی ارتباط بین انواع انرژی تجدیدپذیر و ویژگی‌های انتخاب شده است.

این نمودار به ما کمک می‌کند تا روابط پیچیده بین انواع انرژی تجدیدپذیر و ویژگی‌های مهم مانند مشاغل ایجاد شده و تولید انرژی را به صورت بصری مشاهده کنیم. با مشاهده این نوع نمودار می‌توانیم الگوهایی را شناسایی کنیم که ممکن است در داده‌ها مخفی باشند. در اینجا تعداد شغل‌هایی که هر نوع انرژی ایجاد کرده روی خط وصل شده به job Energy created مشخص است. همچنین مقدار انرژی تولید شده به مگاوات در ساعت روی خطوط متصل به Production نوشته شده‌اند.



نمودار دایره‌ای

۴-۵



یکی دیگر از روش‌هایی که می‌توان تعداد شغل‌های ایجاد شده توسط هر نوع انرژی را با هم مقایسه کرد، استفاده از نمودار دایره‌ای (Pie Chart) است. میزان درصد شغل‌هایی که به هر نوع انرژی اختصاص دارند روی شکل مشخص شده است. مشاهده می‌شود انرژی بادی با اختلاف کمی بیشترین تعداد شغل را به خود اختصاص داده است و ۱۴,۷ درصد از کل را تشکیل می‌دهد. همچنین انرژی موج با ۱۳,۹ درصد کمترین تعداد شغل را ایجاد کرده است. در کل نتایج به هم نزدیک هستند و اختلاف آنچنانی وجود ندارد.

۶. کاربرد نتایج در تصمیم‌گیری

تصمیمات عملی:

تحلیل داده‌های مربوط به منابع انرژی تجدیدپذیر می‌تواند به تصمیمات عملی و توصیه‌های سیاستی منجر شود. این تصمیمات شامل موارد زیر است:

- توسعه پروژه‌های انرژی تجدیدپذیر: شناسایی انواع منابع انرژی با بازدهی بالا و تأثیر مثبت بر کاهش آلودگی و پیشنهاد سرمایه‌گذاری بیشتر در این منابع.
- افزایش ظرفیت ذخیره‌سازی انرژی: با توجه به تحلیل کارایی سیستم‌های ذخیره‌سازی، می‌توان برنامه‌ریزی برای ارتقاء این سیستم‌ها و افزایش ظرفیت ذخیره‌سازی انرژی را پیشنهاد داد.
- بهینه‌سازی مشوق‌های مالی: تحلیل مشوق‌های مالی و تأثیر آنها بر توسعه پروژه‌ها می‌تواند به بهینه‌سازی این مشوق‌ها کمک کند تا بهره‌وری بیشتری از منابع مالی حاصل شود.
- ارتقاء یکپارچگی شبکه: با شناسایی سطح یکپارچگی پروژه‌ها با شبکه برق، می‌توان راهکارهایی برای افزایش یکپارچگی و بهبود عملکرد شبکه ارائه داد.

درک تأثیر:

- بررسی تأثیر بالقوه و امکان‌سنجی اقدامات پیشنهادی در سناریوهای واقعی از اهمیت بسیاری برخوردار است:
- تحلیل تأثیرات زیست‌محیطی: کاهش انتشار گازهای گلخانه‌ای و آلودگی هوا از اهداف اصلی پروژه‌های انرژی تجدیدپذیر است. تحلیل داده‌ها نشان می‌دهد که توسعه این پروژه‌ها می‌تواند به بهبود کیفیت هوا و کاهش تغییرات اقلیمی کمک کند.
 - اشتغال‌زایی: با توجه به تحلیل تعداد مشاغل ایجاد شده توسط پروژه‌ها، می‌توان به اهمیت اقتصادی این پروژه‌ها و تأثیرات مثبت آنها بر اشتغال‌زایی و بهبود وضعیت اقتصادی جوامع پی برد.
 - پایداری اقتصادی: بررسی هزینه‌های سرمایه‌گذاری اولیه و منابع تأمین مالی می‌تواند به تصمیم‌گیری‌های اقتصادی پایدارتر کمک کند. با شناخت منابع مالی مؤثر و بهینه‌سازی هزینه‌ها، می‌توان به توسعه پایدار انرژی تجدیدپذیر دست یافت.

بهبود سیاست‌ها و راهکارها:

نتایج این تحلیل‌ها می‌تواند به بهبود سیاست‌ها و راهکارهای اجرایی کمک کند:

- **تدوین سیاست‌های حمایتی**: بر اساس داده‌های مشوق‌های مالی و تأثیرات آنها، می‌توان سیاست‌های حمایتی جدیدی تدوین کرد که به توسعه بیشتر انرژی‌های تجدیدپذیر کمک کند.
- **پیشنهادات سیاستی برای کاهش آلودگی**: با توجه به شاخص‌های کاهش آلودگی هوا و کاهش انتشار گازهای گلخانه‌ای، می‌توان پیشنهادات سیاستی جدیدی برای کنترل و کاهش آلودگی‌ها ارائه داد.
- **افزایش شفافیت و آگاهی**: تحلیل داده‌ها می‌تواند به افزایش شفافیت در حوزه انرژی تجدیدپذیر و افزایش آگاهی عمومی کمک کند. با ارائه گزارش‌های جامع و دقیق، می‌توان به جلب حمایت عمومی و افزایش مشارکت مردم در پروژه‌های انرژی تجدیدپذیر پرداخت.

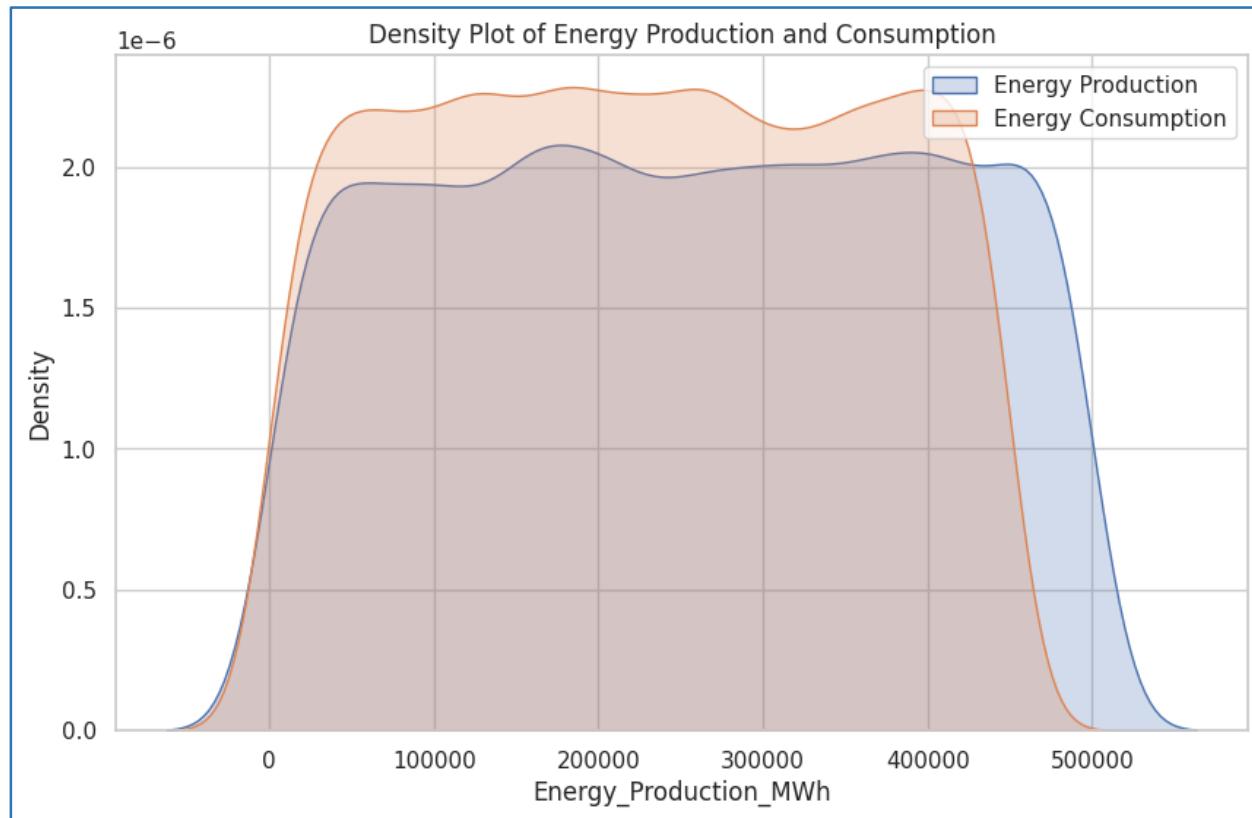
این تحلیل داده‌ها به ما امکان می‌دهد تا با دیدی جامع‌تر و دقیق‌تر به ارزیابی و بهبود منابع انرژی تجدیدپذیر بپردازیم. نتایج حاصل از این تحلیل‌ها می‌تواند به سیاست‌گذاران، مدیران و سرمایه‌گذاران کمک کند تا تصمیمات بهتری در زمینه توسعه و بهره‌برداری از منابع انرژی تجدیدپذیر اتخاذ کنند. با توجه به اهمیت انرژی‌های تجدیدپذیر در کاهش آلودگی و تأمین انرژی پایدار، این تحلیل‌ها نقش حیاتی در دستیابی به اهداف زیستمحیطی و اقتصادی ایفا می‌کنند.

۷. محدودیت‌ها و پیشنهادها

همانطور که در طول پژوهش مشاهده شد همواره با چالش‌ها و محدودیت‌هایی در این تحقیق مواجه شدیم که برخی از آن‌ها رفع شد. اما مواردی از آن‌ها بی‌پاسخ مانده که پیشنهاد می‌شود در پژوهش‌های بعدی مورد بررسی قرار بگیرند.

تشخیص توزیع دقیق داده‌ها:

در طی بررسی‌های انجام شده دیده شد که هیچکدام از متغیرها از توزیع نرمال یا نمایی پیروی نمی‌کنند اما همانطور که در نمودار چگالی زیر مشاهده می‌شود، با توجه به شب تند دم راست و چپ، داده‌ها مشکوک به توزیع بتا می‌باشند و نیاز است با روش‌های آماری موجود این موضوع بررسی شود.



نرمال سازی داده‌ها:

در این پژوهش، پژوهشگران از روش‌های در دسترس از جمله تبدیل لگاریتمی، معیار استاندارد و حداکثر مقدار مطلق برای نرمال کردن داده‌ها استفاده کردند که هیچکدام نتیجه موفقیت‌آمیزی نداشتند. پیشنهاد می‌شود در ادامه تشخیص توزیع متغیرها، بررسی روش‌های مختلف برای نرمال کردن متغیرها هم مورد بررسی قرار بگیرند.

لازم به ذکر است که در استفاده از آزمون‌های نرمالیتی محدودیت‌هایی در استفاده از آزمون شاپیرو و کولموگروف اسمیرنوف بود. در این دو آزمون به دلیل بالا بودن حجم مشاهدات نتایج آزمون دقیق و مورد اعتماد نبود و پیشنهاد می‌شود در صورت نیاز برای بررسی این مسئله از آزمون جارکیو-برا استفاده شود که به حجم داده‌ها حساسیت کمتری دارد.

همبستگی متغیرها:

محدودیت دیگری که در این مجموعه داده وجود دارد عدم وجود رابطه خطی بین متغیرها است. از راه‌های پیشنهادی برای بررسی این موضوع می‌توان به نمونه‌گیری از داده‌ها و بررسی وجود رابطه خطی بین نمونه‌ها است. راه دیگر پیشنهادی بررسی وجود رابطه غیرخطی همچن رابطه چندجمله‌ای، سینوسی و کسینوسی، توانی و... بین متغیرها می‌باشد.