* اهمیت اصلی استفاده از داده ‌های حجیم تبدیل آن‌ها به داده ‌های قابل ‌فهم و باارزش برای استفاده در پروژه و رسیدن به اهداف طرح است.
* با توجه به حجم بزرگ داده ‌های حجیم و غیر ساختاریافته، نیاز به منابع مالی، زمانی و افراد خبره برای پردازش و رسیدن به داده ‌های باارزش وجود دارد.
* -تحلیل داده ‌ها به ۴ دسته تقسیم می‌شود:

 1) \*\*تحلیل توصیفی\*\*: پاسخ به سوال "چه چیزی اتفاق افتاده است؟" و با استفاده از آنالیز داده ‌های گذشته به یک گزارش از داده ‌ها می‌توان دست یافت.

 2) \*\*تحلیل تشخیصی\*\*: پاسخ به سوال "چرا یک اتفاق خاص رخ داده است؟" که بیشتر بر همبستگی تمرکز دارد.

 3) \*\*تحلیل پیش‌بینی ‌کننده\*\*: پیش‌بینی وقایع و پیامدهای آینده بر اساس داده ‌های گذشته و کنونی که بیشتر از آنالیزهای رگرسیونی و مدل ‌های پیش‌بینی ‌کننده یادگیری ماشین استفاده می‌شود.

 4) \*\*تحلیل تجویزی\*\*: پس از انجام همه مراحل، تصمیماتی که باید گرفته شود و پیشنهاداتی برای اقداماتی که باید انجام شود، در نظر گرفته می‌شود.

* استدلال ‌های مورد استفاده به ۳ دسته تقسیم می‌شوند: استدلال استنتاجی، استقرایی، ابداعی.
* کیفیت داده ‌های مورد نیاز برای انجام پروژه ‌ها از کمیت آن‌ها اهمیت بیشتری دارد و باید ویژ گی‌های مهم و مؤثر برای هر طرح با توجه به اهداف پروژه مشخص شوند.
* از چالش ‌های داده ‌های حجیم می‌توان به موارد زیر اشاره کرد: رشد داده ‌ها، زیرساخت، نیروی متخصص، منابع داده ‌ها، کیفیت داده ‌ها، حفظ امنیت و حریم خصوصی داده ‌ها، مقاومت افراد ذی‌نفع و سیاست‌گذاران، تکه ‌تکه بودن داده ‌ها، نبود استراتژی داده، سیاست ‌گذاری برای حاکمیت داده ‌ها، تصویرسازی، مباحث اخلاقی.
* نیاز به یک سیستم مدیریت و حاکمیت داده در هر سازمانی وجود دارد که وظایف و مراحل انجام هر مرحله را برای هر سازمان بر اساس گایدلاین ‌های موجود برای استانداردسازی و مقررات عمومی حفاظت از داده ‌ها تعیین و اجرا کند.
* برای استقرار و گسترش طرح ‌ها چند نکته باید در نظر گرفته شود:

 1) هدف انجام پروژه مشخص شود.

 2) تاثیر پروژه در سازمان تعیین شود.

 3) بر اساس آن زیرساخت ‌ها، منابع، شاخص ‌های کلیدی مشخص شوند.

 4) مهارت ‌ها و فناوری ‌ها و ابزارهای مورد نیاز فراهم شوند.

 5) داده ‌ها مشخص، جمع ‌آوری و آماده ‌سازی شوند.

 6) نوع تحلیل ‌ها و مدل‌های مورد استفاده مشخص شوند.

 7) ساختارهای پایگاه داده، ذخیره ‌سازی، دسترسی و استفاده ‌های آینده مشخص شوند.

 8) طرح اجرا شود.

 9) شکاف ‌های موجود بین پروژه اخیر و توانمن دی‌های آینده در نظر گرفته شود تا امکان گسترش طرح در آینده وجود داشته باشد.

 10) در محیط آزمایشی طرح تست شود و سپس نتایج گزارش شوند.

 11) طرح به مرحله تجاری ‌سازی و استفاده در صنعت در مقیاس بزرگ‌تر برسد.