



# डाटा विश्लेषण पर दिशानिर्देश

भारत के नियंत्रक-महालेखापरीक्षक का कार्यालय

2017

## विषय-सूची

प्राक्कथन .....	
1. डाटा विश्लेषण.....	1
2. डाटा अधिग्रहण और तैयारी .....	5
3. डाटा विश्लेषण एवं मॉडलिंग.....	18
4. लेखापरीक्षा में डाटा विश्लेषण का उपयोग.....	32
अनुबंध .....	38

## प्राक्कथन

सार्वजनिक वस्तुओं और सेवाओं के वितरण को बढ़ाने के लिए वर्तमान शासन प्रणाली में प्रौद्योगिकी महत्वपूर्ण भूमिका निभाती है। विविध प्रौद्योगिकी प्रणालियां डाटा विश्लेषणों के लिए अनगिनत अवसर उत्पन्न कर विभिन्न प्रकारों से निरंतर डाटा उत्पन्न कर रही हैं।

प्रतिक्रियाशील उच्चतम लेखापरीक्षा संस्थान के रूप में, हमें इस प्रकार की प्रगति के साथ तालमेल रखने के लिए संस्थागत रूप से स्फूर्त होना होगा तथा डाटा विश्लेषण में विकसित हो रहे अवसरों को स्वीकार करना होगा। 2015 में बनाई गई बिग डाटा प्रबंधन नीति में विभाग के लिए डाटा विश्लेषण फ्रेमवर्क की विस्तृत रूपरेखा की कल्पना की गई। डाटा प्रबंधन और विश्लेषण के लिए केंद्र का निर्माण इस फ्रेमवर्क की स्थापना का पहला कदम था।

डाटा विश्लेषण के दिशा-निर्देश पद्धति को संस्थागत करने और विभाग में डाटा विश्लेषण का प्रयोग करने की ओर एक महत्वपूर्ण कदम है। ये दिशानिर्देश डाटा विश्लेषण की अवधारणा की व्याख्या करते हैं, डाटा विश्लेषण की प्रक्रिया की रूपरेखा बनाते हैं और डाटा विश्लेषण मॉडलों के विकास पर विचार करते हैं। डाटा विश्लेषण एक विकासशील विधा है और इसलिए इन दिशानिर्देशों को आवधिक समीक्षित और अद्यतित किया जाना होगा।

मुझे विश्वास है कि विभाग के अधिकारी और स्टाफ इन दिशानिर्देशों को उपयोगी पाएंगे और लोक लेखांकन और लेखापरीक्षण की गुणवत्ता को बढ़ाने की दिशा में अर्थपूर्ण रूप से उनका प्रयोग करेंगे।



शशि कान्त शर्मा

भारत के नियंत्रक-महालेखापरीक्षक

सितम्बर 2017

## 1. डाटा विश्लेषण

### प्रस्तावना

1.1 डाटा विश्लेषण डाटा से ज्ञान (पूरी जानकारी) अर्जित करने के लिए डाटा विज्ञान<sup>1</sup> पद्धतियों का अनुप्रयोग है। इसमें डाटा एकत्रित करने, डाटा तैयार करने से आरंभ करते हुए तथा उसके बाद प्रासंगिक परिज्ञान प्राप्त करने के लिए विभिन्न डाटा विश्लेषण तकनीक लागू करने तक कई सोपान शामिल होते हैं। परिज्ञान में ट्रेंड, पैटर्न, विचलन, विसंगतियों तथा विश्लेषण, मॉडलिंग या विजुलाइज़ेशन के माध्यम से पहचाने गये डाटा तत्वों में संबंध, शामिल होते हैं, जिनका प्रयोग नियोजन और लेखापरीक्षण में किया जा सकता है, शामिल होते हैं लेकिन ये इन तक सीमित नहीं होते।

डाटा विश्लेषण जानकारी आधारित निर्णय लेने को सशक्त बनाने में प्रतियोगी लाभ देता है। चूँकि यह एक उभरती हुई विधा है, इसलिए दोनों लोक तथा निजी क्षेत्र में इसके संभव प्रयोग अभी भी प्रयोगात्मक तथा अन्वेषण के अंतर्गत हैं।

1.2 ये दिशानिर्देश भारतीय लेखापरीक्षा तथा लेखा विभाग (आईएएंडएडी) के लेखापरीक्षण कार्य में डाटा विश्लेषण लागू करने की कार्यप्रणाली को निर्धारित करते हैं। डाटा विश्लेषण सिद्धांत और विधियां, तथापि, लेखाकरण और प्रशासन के कार्यक्षेत्र में लागू होंगी।

1.3 ये दिशानिर्देश सितम्बर 2015 में जारी बिग डाटा मैनेजमेंट पॉलिसी के अनुसरण में जारी किए गए हैं तथा आईएएंडएडी, विशेषकर लेखापरीक्षा में, डाटा विश्लेषण के प्रयोग में उत्तरगामी कदम उठाने होंगे। ये दिशानिर्देश वर्तमान निष्पादन लेखापरीक्षण, अनुपालन लेखापरीक्षण, वित्तीय लेखापरीक्षण, लेखापरीक्षण मानकों और आईएएंडएडी में अन्य संबंधित निर्देशों तथा नियम पुस्तकों पर वर्तमान दिशानिर्देशों पर प्रभावी होंगे।

---

<sup>1</sup> डाटा विज्ञान का अभिप्राय जानकारी के विशाल संग्रह एकत्रीकरण, तैयारी, विश्लेषण, दूरदृष्टि, प्रबंधन तथा संरक्षण के संबंधी उभरते हुए कार्यक्षेत्र से है। यद्यपि डाटा विज्ञान का संबंध डाटाबेस और कम्प्यूटर विज्ञान कई विभिन्न प्रकार के कौशल जैसे क्षेत्रों से बहुत घनिष्ठ है - जिसमें गणितइतर कौशल आवश्यक है - की आवश्यक होती है - एन इंट्रोडक्शन टू डाटा साइंस - जैफरी स्टेनटोन - साईराक्स यूनिवर्सिटी।

## एकल पहल और पेशेवर निर्णय के कार्यक्षेत्र

1.4 जबकि ये दिशानिर्देश आदेशात्मक प्रकृति के हैं, लेकिन ये महालेखाकार<sup>2</sup> के पेशेवर निर्णयों का स्थान लेने को अभिप्रेत नहीं हैं। महालेखाकार से इन दिशानिर्देशों में निर्धारित प्रावधानों के लिए स्थिति या विषय आधारित समायोजन करने की अपेक्षा की जाती है। तथापि, महालेखाकार से इन दिशानिर्देशों से प्राप्त होने वाले महत्वपूर्ण उपक्रमों का तर्काधार प्रलेखित करने तथा सक्षम प्राधिकारी से प्राधिकार प्राप्त करने की अपेक्षा की जाएगी।

## डाटा विश्लेषण और आईएण्डएडी

1.5 आईएण्डएडी का लेखापरीक्षा अधिदेश बहुत विस्तृत है, जिसमें संघ सरकार और राज्य सरकार के साथ-साथ निकाय या प्राधिकरणों जैसे कि सांविधिक निगम, सरकारी कम्पनियां, सोसाइटी के रूप में गठित स्वायत्त निकाय, न्यास या अलाभकारी कम्पनियां, शहरी और स्थानीय निकाय तथा ऐसे कोई भी अन्य निकाय या प्राधिकरण जिनके लेखापरीक्षण का दायित्व भारत के नियंत्रक-महालेखापरीक्षक को सौंपा गया हो, शामिल है। लेखापरीक्षण योग्य स्वत्व के लेखाओं, वाउचर और रिपोर्ट जिस प्रकार लेखापरीक्षा कार्यालय और/या लेखा कार्यालय में प्राप्त हुए हो और इनमें ऑनलाईन डाटा, जानकारी और दस्तावेज़ शामिल हो सकते हैं, के संदर्भ में लेखापरीक्षा संचालित की जाती है। लेखापरीक्षण मानकों में लेखापरीक्षा के निर्णय के समर्थन के लिए पर्याप्त और समुचित प्रमाण प्राप्त करने के साथ-साथ लेखापरीक्षा के अंतर्गत संस्था, कार्यक्रम, कार्यकलाप या कार्यान्वयन संबंधी निष्कर्ष भी एकत्रित करना परिकल्पित है। सीमित उपलब्ध संसाधनों के साथ, लेखापरीक्षा जोखिम आधारित पद्धति अपनाती है और विश्लेषणात्मक प्रक्रिया, नियंत्रणों की जांच और लेखापरीक्षा की योजना और कार्यान्वयन के दौरान उपलब्ध और चयनित डाटा पर नियंत्रण और अनुगामी जांच बिंदुओं की जांच करती है।

तीव्र कम्प्यूटराईज़ेशन के साथ, लेखापरीक्षा योग्य सत्त्वों के अधिकतर कार्यकलाप विभिन्न आईटी प्रणालियों में इलेक्ट्रॉनिकली रिकॉर्ड किए जाते हैं। इन इलेक्ट्रॉनिक रिकॉर्ड या 'डाटा' का यदि निर्वचन उचित प्रकार से हो, तो पिछली घटनाओं पर जानकारी प्रदान कर सकता है, वर्तमान में सुधारात्मक कार्रवाई पर निर्देश दे सकता है

---

<sup>2</sup> महालेखाकार पद का अभिप्राय आईएण्डएडी में वरिष्ठ प्रशासनिक ग्रेड की रैंक के विभागाध्यक्ष (एचओडी) और उपर से है।

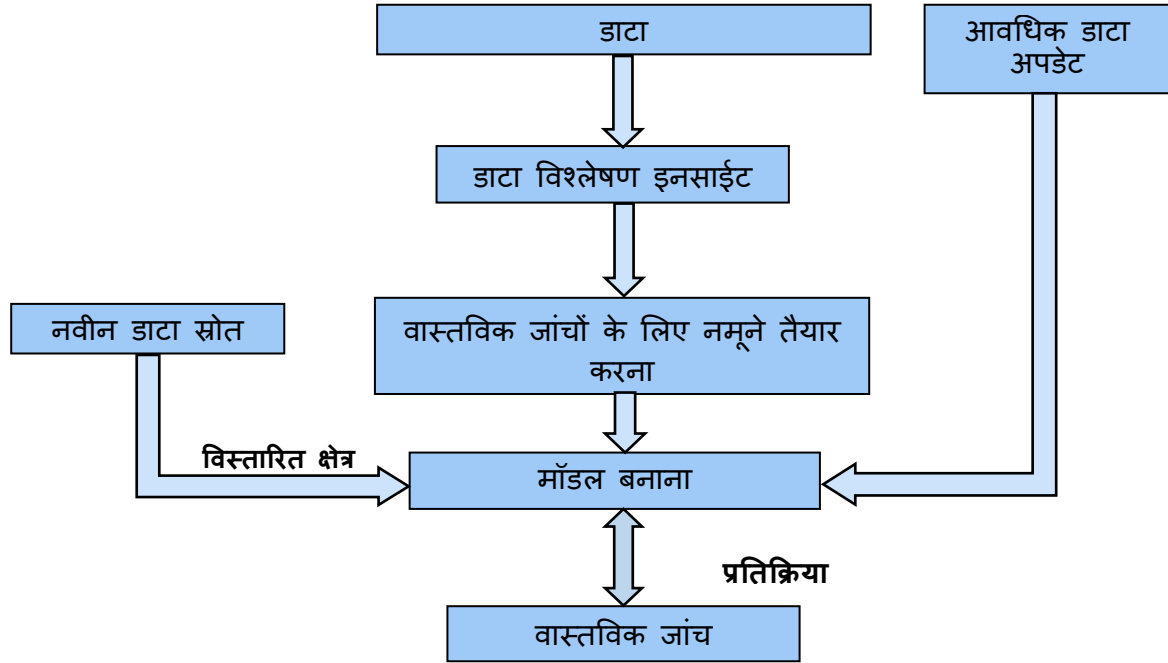
और लेखापरीक्षक की क्षमता में वर्धन करते हुये भविष्य की घटनाओं का अंदाजा लगा सकता है।

- 1.6 डाटा विभिन्न प्रकारों और विभिन्न स्रोतों से लेखापरीक्षा के लिए उपलब्ध है। डाटा विश्लेषण इन डाटा सेटों को विश्लेषित करने के लिए क्षमता प्रदान करता है और डाटा सेटों और उनके वैरिएबल में पैटर्न, ट्रेंड, विवरण, अपवाद, अनियमितता और संबंधों की पहचान करके सभी जानकारी प्राप्त करने के द्वारा लेखापरीक्षा प्रक्रिया में मदद करता है। इस प्रकार आहरित सभी जानकारी प्राथमिक रूप से रूचि और जोखिम के क्षेत्रों और अपवादों की पहचान के द्वारा लेखापरीक्षा की दिशा निर्धारित करने में सहायता करेगी।

### लेखापरीक्षा में डाटा विश्लेषण

- 1.7 डाटा विश्लेषण का आरंभ किसी लेखापरीक्षा विशेष के लिए विभिन्न डाटा स्रोतों की पहचान करने और एकत्रीकरण से होता है। विभिन्न डाटा विश्लेषण तकनीकों के माध्यम से डाटा का विश्लेषण लेखापरीक्षित सत्व की कार्यप्रणाली पर अंतर्दृष्टि प्रदान करता है। इस प्रकार के अभ्यास से पहचाने गये जोखिम क्षेत्र या रूचि क्षेत्र लेखापरीक्षा उद्देश्यों की पहचान करने और लेखापरीक्षा डिज़ाइन मैट्रिक्स विकसित करने में सहायता प्रदान करेंगे। डाटा विश्लेषण लेखापरीक्षा इकाइयों के नमूनों जहां अनुवर्ती जांच संचालित की जाएगी, की पहचान करने में सहायता करेगा।
- 1.8 विभिन्न विश्लेषणों को पुनः प्रयोग करने योग्य डाटा विश्लेषण मॉडल के रूप में बनाया जा सकता है। यह सुनिश्चित करेगा कि डाटा विश्लेषण के परिणाम, डाटा के आवधिक अद्यतन के साथ बार-बार प्रयोग किए जा सकते हैं। आवधिक रूप से डाटा प्राप्त करने के लिए तन्त्र स्थापित करना ऐसे अभिगम के लिए निर्णायक होगा। एक बार बनाये गये इस मॉडल का कार्यक्षेत्र अनुवर्ती जांचों से प्राप्त फीडबैक और अतिरिक्त डाटा स्रोतों को शामिल करके बढ़ाया जा सकता है। इस प्रकार, आईएएंडएडी में डाटा विश्लेषण किसी विशिष्ट लेखापरीक्षा के लिए एक बार की प्रक्रिया नहीं है, वरन् यह समय के अनुसार विकसित होती रहेगी।

1.9 प्रक्रिया का योजनाबद्ध चित्र निम्नानुसार है:



चित्र 1- डाटा विश्लेषण प्रक्रिया

अनुवर्ती अध्यायों में डाटा विश्लेषण प्रक्रिया विस्तृत रूप से वर्णित की गई है।

1.10 आईएएंडएडी में डाटा विश्लेषण कार्यकलाप चलाने के लिए सेंटर फॉर डाटा मैनेजमेंट एंड एनलेटिक्स (सीडीएमए) नोडल बॉडी होगी। डाटा विश्लेषण के भावी निर्देशन में डाटा विश्लेषण और उच्च अनुसंधान और विकास पर क्षेत्रीय कार्यालयों को सीडीएमए निर्देशन प्रदान करेगा।

आईएएंडएडी में डाटा विश्लेषण के लिए परिकल्पित संरचना में, अपनी वार्षिक योजना के अनुसार क्षेत्रीय कार्यालय द्वारा डाटा विश्लेषण किया जाना है। क्षेत्रीय कार्यालय में डाटा विश्लेषण कार्यकलाप की जिम्मेदारी कार्यालयाध्यक्ष (एचओडी) की होगी, जो डाटा विश्लेषण समूह का गठन करेगा। ग्रुप अधिकारी के प्रभार के अन्तर्गत क्षेत्रीय कार्यालयों में गठित डाटा विश्लेषण समूह क्षेत्रीय कार्यालय में डाटा विश्लेषण करने के लिए उत्तरदायी होगा। डाटा विश्लेषण से लेखापरीक्षा के लिए अर्थपरक अंतर्दृष्टि प्राप्त करने के लिए, लेखापरीक्षा के क्षेत्र का ज्ञान अनिवार्य है। इसलिए विश्लेषण की प्रक्रिया में डाटा विश्लेषण समूह के तकनीकी ज्ञान और क्षेत्रीय कार्यालयों में क्रियान्वयन समूहों से डॉमेन विशेषज्ञता, एक दूसरे के पूरक के रूप में, समूहगत प्रयास परिकल्पित

है। डाटा विश्लेषण कार्यकलापों के लिए सांकेतिक भूमिका कार्यभार अनुबंध 1 में दिया जाता है।

### बाहरी विशेषज्ञ भर्ती करना

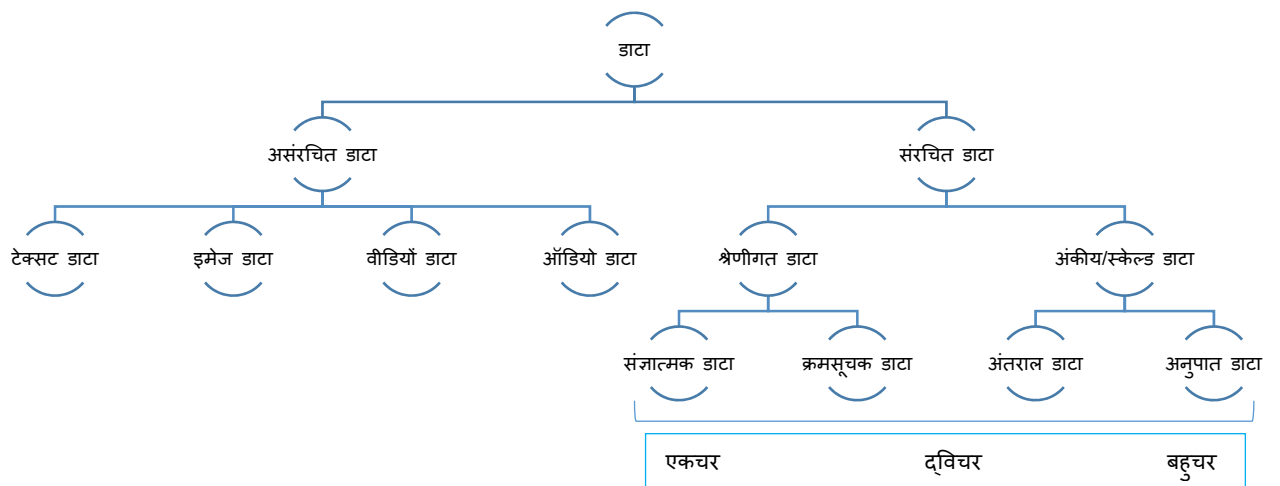
1.11 विशेषज्ञ क्षेत्रों में, क्षेत्रीय कार्यालय बाहरी विशेषज्ञों को भर्ती करने पर विचार कर सकते हैं, यदि ऐसी आवश्यकता औचित्यपूर्ण है। तथापि, बाहरी विशेषज्ञों की भर्ती, समय-समय पर आईएएंडएडी द्वारा जारी दिशानिर्देशों के अनुसार की जानी चाहिए। इस प्रकार की भर्ती के लिए विशिष्ट क्षेत्र डाटा हैडलिंग, एडवांस्ड डाटा विश्लेषण तकनीक या डाटा रिपॉजिटरी प्रबंधन से संबंधित हो सकते हैं।

## 2. डाटा अधिग्रहण और तैयारी

2.1 डाटा विश्लेषण प्रक्रिया में डाटा अधिग्रहण, डाटा तैयार करना, डाटा विश्लेषण, परिणाम और विश्लेषण मॉडल शामिल होते हैं। यह अध्याय आंकड़ों की पहचान और संग्रहण के साथ-साथ एकत्रित आंकड़ों को संभालने और विश्लेषण के लिए तैयार करता है। तथापि, अधिग्रहण, तैयारी और विश्लेषण की प्रक्रिया को आरंभ करने से पूर्ण डाटा के प्रकार और उसके स्रोत को समझाना महत्वपूर्ण है।

### डाटा प्रकारों को समझना

2.2 डाटा विश्लेषण का केन्द्रबिन्दु 'डाटा' है। तथ्यों और कारणों को एक अर्थपूर्ण विवेचना देने के लिए डाटा को मापा, संग्रह, विश्लेषण और देखा जा सकता है। डाटा को निम्न प्रकार से समझा तथा वर्गीकृत किया जा सकता है:



चित्र 2- डाटा के प्रकार



- **असंचरित और संरचित डाटा:** असंचरित डाटा में पाठ्य, प्रतिरूप, ओडियो या विडियो डाटा जैसे डाटा शामिल होते हैं जिन्हें सांख्यिकीय या गणितीय विश्लेषण के लिए तत्काल 'सारणीबद्ध' नहीं किया जा सकता है। दूसरी तरफ संचरित डाटा सारणी रूप से संबंध रखता है। संचरित डाटा को श्रेणीगत या अंकीय किया जा सकता है।
- **श्रेणीगत या अंकीय डाटा:** डाटा नाम मात्र का (डाटा क्रमबद्ध नहीं होता है अर्थात: किसी व्यक्ति का नाम, लिंग) या क्रमानुसार (डाटा क्रमबद्ध होता है अर्थात सेवा की गुणवत्ता पर आधारित रैंक: अत्याधिक सन्तुष्ट, सन्तुष्ट, सन्तुष्ट नहीं) हो सकता है। अंकीय डाटा के उदाहरणों में अन्तराल डाटा (अर्थात तापमान जो कि मानों की विभिन्नताओं को पहचान के योग्य है) या अनुपात डाटा (अर्थात किसी कम्पनी का व्यय जिसकी एक दूसरे के बहुल व्ययों के साथ तुलना की जा सकती है) हो सकते हैं।

प्रचालन	नाम मात्र	क्रमवाचक	मध्यान्तर	अनुपात
काउंट	✓	✓	✓	✓
मूल्यों का क्रम		✓	✓	✓
विधि	✓	✓	✓	✓
औसत		✓	✓	✓
माध्य			✓	✓
जोड़ना/घटाना			✓	✓
गुणा/विभाजन				✓
क्या सत्य शून्य मौजूद है				✓

चित्र 3- डाटा के प्रकार के साथ संभावित परिचालन

- **परिवर्तनियों की संख्या - एकल परिवर्तनीय, द्विचर अथवा बहुल परिवर्तनीय डाटा:** डाटा सेट में परिवर्तनियों की संख्या के आधार पर, इसे एकल परिवर्तनीय, द्विचर अथवा बहुल परिवर्तनीय डाटा कहा जा सकता है। एकल परिवर्तनीय डाटा में केवल एक परिवर्तनीय होता है। यह प्रकृति में विवरणात्मक होता है। एकल परिवर्तनीय डाटा के विश्लेषण में डाटा में स्वरूपों के संक्षिप्तिकरण और पहचान शामिल होती है। द्विचर डाटा में दो परिवर्तनीय होते हैं और दो परिवर्तनियों के संबंध को समझने के लिए सांख्यिकीय विश्लेषण को अपनाया जा सकता है। वे एक्स-वाई आधार पर दर्शाये जा सकते हैं और स्कैटर प्लॉट की भाँति प्लॉटों द्वारा दृश्य नमूना डाटा के इस प्रकार के स्वरूप को समझने के लिए उपयोगी होंगे। बहुल

परिवर्तनीय डाटा में कई परिवर्तनीय शामिल होते हैं। सांख्यिकीय विश्लेषण डाटा विश्लेषण के लिए और परिवर्तनियों के बीच संबंध और निर्भरता को खोजने के लिए अपेक्षित होगा। दृश्य चित्रण विभिन्न परिवर्तनियों के बीच संबंध समझने के लिए एक महत्वपूर्ण साधन है और प्लॉट एक्स, वाई और जैड के तीन आयामों पर दिखाया जा सकता है। इस प्रकार प्लॉटों में उपयुक्त दृश्यात्मक दृष्टिकोण अपनाते हुए तीन परिवर्तनियों से अधिक को शामिल किया जा सकता है।

## डाटा के स्रोत

2.3 आईएएंडएडी को उपलब्ध डाटा के विभिन्न स्रोतों की पहचान डाटा प्रबंधन प्रारूप का आधार बिंदु है। विस्तृत डाटा प्रबंधन नीति विभिन्न डाटा स्रोतों को निम्नवत श्रेणीबद्ध करती है:

**आंतरिक डाटा स्रोत:** इसमें शामिल हैं

- सम्मिलित वित्त और राजस्व लेखा
- वीएलसी डाटा बेस
- लेखा व हक. कार्यालयों में जीपीएफ और पेंशन डाटा
- लेखापरीक्षा प्रक्रिया के माध्यम से उत्पन्न डाटा
- विभाग में उपलब्ध अन्य डाटा

**बाह्य डाटा स्रोत:** इसमें शामिल हैं

क) अपनी व्यवसायिक क्षमता में विभाग के पास उपलब्ध **लेखापरीक्षित सत्त्वों का डाटा जिसमें शामिल है**

- लेखापरीक्षित सत्त्वों का वित्तीय और गैर वित्तीय डाटा
- लाभभोगी डाटा बेस सहित कार्यक्रम निर्दिष्ट डाटा
- लेखापरीक्षित सत्त्वों से संबंधित अन्य डाटा

ख) **तीसरी पार्टी डाटा** जिसमें पब्लिक डोमेन में उपलब्ध डाटा समाविष्ट होता है और शामिल करता है:

- सरकारी और सांविधिक प्राधिकरणों द्वारा प्रकाशित डाटा जैसे
  - जनगणना डाटा
  - एनएसएसओ डाटा
  - विभिन्न मंत्रालयों/विभागों द्वारा प्रकाशित डाटा

- data.gov.in में उपलब्ध डाटा
- विभिन्न कमिश्नों की रिपोर्ट
- संघ सरकार/राज्यों से संबंधित अन्य रिपोर्ट और डाटा
- पब्लिक डोमेन में उपलब्ध अन्य डाटा
  - एनजीओ द्वारा प्रकाशित सर्वेक्षण और सूचना
  - सीआईआई, फिक्की/नेस्कॉम आदि द्वारा प्रकाशित औद्योगिक विशेष सूचना
  - विभिन्न संगठनों द्वारा प्रकाशित क्षेत्र विशेष सूचना
  - सामाजिक मीडिया आदि।

2.4 क्षेत्रीय कार्यालयों में यह स्थिति पैदा हो सकती है जहां अपेक्षित डाटा मैनुअल रूप में उपलब्ध है। क्षेत्रीय कार्यालयों को तब यह निर्णय करना चाहिए कि क्या मैनुअल डाटा को इलेक्ट्रॉनिक डाटा सेट बनाते हुए इन्हें इलेक्ट्रॉनिक रूप में बदला जा सकता है। उदाहरणतः लेखापरीक्षा कार्यालयों में प्राप्त संस्वीकृति आदेशों में दिये गये विवरण इलेक्ट्रॉनिक डाटा फाईल में बदले जा सकते हैं, जिन्हें डाटा विश्लेषणों के लिए प्रयोग किया जा सकता है।

### डाटा की पहचान करना

2.5 लेखापरीक्षा योग्य सत्त्वों, पर एक विस्तृत डाटा के संग्रहण और अनुरक्षण के भाग के रूप में, क्षेत्रीय कार्यालयों को अपने क्षेत्राधिकार के अंतर्गत लेखापरीक्षित सत्त्वों/तीसरी पार्टी डाटा के पास इलेक्ट्रॉनिक डाटा की उपलब्धता की पहचान करने के लिए एक तंत्र तैयार करना चाहिए और आवधिक रूप से उन्हें अद्यतित करना चाहिए।

### डाटा अधिग्रहण

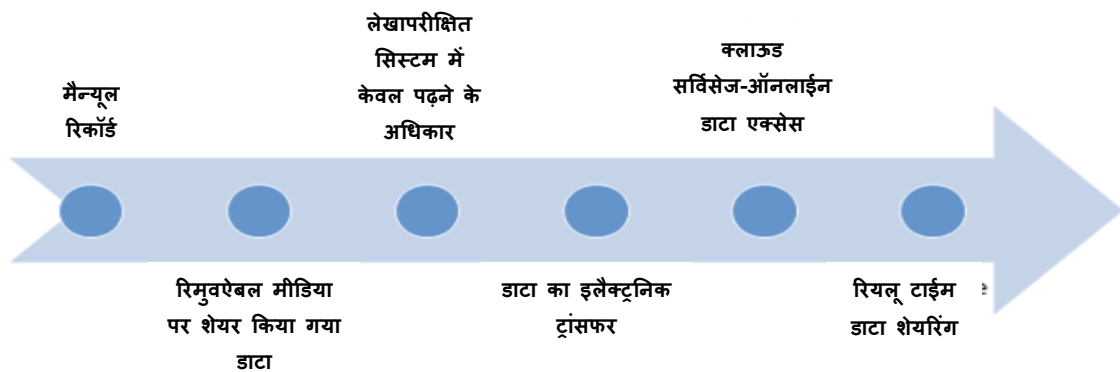
2.6 डाटा अधिग्रहण में संग्रहीत डाटा के स्वामित्व, सुरक्षा और विश्वसनीयता के मद्देनजर डाटा तक पहुँच और संग्रहण शामिल होते हैं।

### डाटा एक्सेस

2.7 चूँकि आईएण्डएडी डाटा विश्लेषणों के लिए अपेक्षित विभिन्न डाटा स्रोतों का मालिक नहीं है, इसलिए डाटा उपलब्धता मध्यम अवधि में एक चुनौती के रूप में ली जाएगी। इस समस्या को कटु बनाना अपने डाटा को विभाजित करना कई लेखापरीक्षित सत्त्वों की अरुचि है। नियंत्रक-महालेखापरीक्षक के कर्तव्य, शक्तियां और सेवा की शर्तें,

अधिनियम 1971 और लेखा और लेखापरीक्षा 2007 विनियमन के प्रासंगिक प्रावधानों से सहायता लेते हुए लेखापरीक्षित सत्त्वों का लगातार प्रतिपादन और मॉनीटरिंग इस मामले को सुलझाने का तरीका होगा।

- 2.8 सिस्टम तक अभिगमन द्वारा सत्त्वों की साईट पर लेखापरीक्षकों को डाटा उपलब्ध कराया जाय। बिना किसी आदान प्रदान के अधिकार के अंतर्गत यह केवल पढ़ने के लिए अभिगमन हो सकता है ताकि सिस्टम का निष्पादन प्रभावित न हो। यह डाटा सत्त्व के वातावरण में तैयार की गई फाईलों के बैकअप द्वारा उपलब्ध कराया जाय और लेखापरीक्षकों के साथ हटाने योग्य मीडिया पर शेयर किया जाय। डाटा को नैटवर्क-लैन या वैन या इंटरनेट या एक वीपीएन, जैसा भी मामला हो, के माध्यम से इलैक्ट्रॉनिक स्थानांतरण उपयोग करते हुए इलैक्ट्रॉनिक रूप से भी शेयर किया जाय।
- 2.9 नीचे दर्शाया गया प्रगामी रास्ता है जिसमें लेखापरीक्षक मैन्यूअल रिकॉर्डों से आरंभ करके ऑनलाइन तक रियल टाइम डाटा शेयरिंग अपने लेखापरीक्षित सत्त्वों से ऑनलाइन डाटा तक पहुँच कर सकता है। तथापि, यह आवश्यक नहीं है कि प्रगति क्रमवार हो और लेखापरीक्षक केवल मैन्यूअल रिकॉर्डों तक पहुँच माध्यमिक चरणों तक जाये बिना रियल टाइम डाटा इलैक्ट्रॉनिक रूप से एक्सेस करना आरंभ कर सकते हैं। डाटा तक एक्सेस केवल लेखापरीक्षकों की क्षमता, लेखाकरण वातावरण और दोनों के बीच बने एक्सेस स्तर पर निर्भर करता है।



चित्र 4- डाटा अभिगमन

- 2.10 डाटा अभिगमन से निपटने का एक तरीका आईटी सिस्टम की डिजाइन स्टेज से लेखापरीक्षा का शामिल होना है जब भी सिस्टम डिजाइन में लेखापरीक्षा की डाटा आवश्यकताओं को सम्मिलित किया जाना संभव हो। यह आवश्यक फार्मेट में डाटा का अधिग्रहण प्रदान करेगा। यह सुनिश्चित करने के लिए, क्षेत्रीय कार्यालय को महत्वपूर्ण

सिस्टम कार्यों के स्तर पर संबंधित सत्वों की लेखापरीक्षा के लिए डाटा आवश्यकताओं को पूरा करने की आवश्यकता होगी जिससे अपेक्षित डाटा तक एक्सेस किया जा सके जब सिस्टम क्रियाशील हो। ये डाटा आवश्यकताएं अपेक्षित सूचना के सेट, डाटा फार्मेट, ट्रांसफर ढंग और लेखापरीक्षा को उपलब्ध कराये जाने वाले डाटा की आवधिकता को कवर कर सकती थी। इसी समय पर, संपूर्ण सिस्टम या संपूर्ण डाटा के एक्सेस, यदि किसी विशिष्ट लेखापरीक्षा के लिए अपेक्षित हैं ऐसे निष्पादन लेखापरीक्षा, सिस्टम लेखापरीक्षा, आईटी लेखापरीक्षा, विशेष लेखापरीक्षा आदि को सिस्टम विकास स्तर पर लेखापरीक्षकों का शामिल कर अलग नहीं करना चाहिए।

### **डाटा एक्सेस मोड के विभिन्न स्तरों पर डाटा हैंडलिंग**

- 2.11 जब डाटा हटाये जाने योग्य मीडिया में शेयर किया जाता है, लेखापरीक्षकों को मीडिया - सीडी, डीवीडी, टेप ड्राइव या एक यूएसबी ड्राइव आदि को चलाने के लिए हार्डवेयर की आवश्यकता होती है। मीडिया चलाने की क्षमता के साथ लेखापरीक्षकों के पास उपयुक्त प्रचालन सिस्टम और डाटा बेस एप्लीकेशन (जैसे आरडीबीएमएस) की आवश्यकता है जहां डाटा को मीडिया से पढ़ा जा सकता है। इस प्रकार, जहां से डाटा प्राप्त किया गया है स्रोत के समान ही वातावरण डाटा को पढ़ने के योग्य बनाया जाता है। रीड ऑनली अधिकार सत्व के सिस्टम के लेखापरीक्षकों को प्रदान किये गये विशिष्ट दृश्य अधिकार हैं जिन्हें अपेक्षित डाटा को देखने/कॉपी करने के अधिकार दिये जाते हैं। डाटा के इलैक्ट्रॉनिक ट्रांसफर में, फाईल फॉर्म में डाटा नेटवर्क जैसे मेल, फाईल ट्रांसफर प्रोटोकॉल आदि द्वारा प्रयोग कर ट्रांसफर किया जाता है। ऑन-लाइन एक्सेस में, डाटा एक रिमोट सर्वर से क्लाउड द्वारा उपलब्ध कराया जाता है। रियल टाइम सिस्टम लाईव सिस्टम तक एक्सेस करवाता है और उसमें उपलब्ध सूचना रियल टाइम मोड में होती है। रियल टाइम डाटा एक्सेस रियल टाइम प्रोसेसिंग की संभावना बनाता है जिससे यह अंतःस्थापित लेखापरीक्षा मॉड्यूल<sup>3</sup> द्वारा लगातार लेखापरीक्षण दृष्टिकोण के विकास के योग्य बन जाता है।

सभी क्षेत्रीय कार्यालयों को डाटा स्रोत संगठनों के साथ उपयुक्त डाटा एक्सेस तंत्र तैयार करने के लिए प्रयास करना चाहिए ताकि उनके डाटा कोष/डाटा विश्लेषण मॉडलों में आवधिक/रियल टाइम आधार पर डाटा तक एक्सेस किया जा सके।

---

<sup>3</sup> अंतःस्थापित लेखापरीक्षा मॉड्यूल- इस प्रकार आईटी सिस्टम के साथ अंतःस्थापित/एकीकृत लेखापरीक्षा मॉड्यूल रियल टाइम डाटा सहित ऑन-लाइन प्राप्त करता है।

## डाटा का संग्रहण

- 2.12 डाटा संग्रहण रूचि के क्षेत्र का एक पूर्ण और सटीक विवरण प्राप्त करने के लिए विभिन्न स्रोत से सूचना एकत्र और मापने का सुव्यवस्थित दृष्टिकोण है। आईटी सिस्टम का डाटा संग्रहण के समय अध्ययन किया जाना चाहिए और समझ जाना चाहिए जो प्रासंगिक डाटा की पहचान और आवश्यकता दर्शाएगा। ये पूर्व डाटाबेस, डाटाबेस से चयनित तालिकाएं, किसी विशेष अवधि स्थिति, वर्ग आदि हेतु विशेष मानदंड/शर्त से संबंधित डाटाबेस या डाटा में तालिकाओं के चयनित डाटा क्षेत्र हो सकते हैं। डाटा आकार के आधार पर, इसे प्लैट फाईल या डंप फाईल फार्मेट में प्राप्त किया जाय। जहां विश्लेषण हेतु प्रासंगिक डाटा/तालिकाओं को प्राप्त करना संभव नहीं है, संपूर्ण डाटा को संग्रहीत किया जाय।
- 2.13 डाटा संग्रहण के समय पर, डाटा की विश्वसनीयता, समग्रता, प्रासंगिकता, उपयोगिता और डाटा सेट की सुरक्षा सुनिश्चित<sup>4</sup> की जानी चाहिए। डाटा (अर्थात-जो डाटा खोया नहीं है) की समग्रता सुनिश्चित करने के लिए, जांच जैसे रिकॉर्ड की कुल संख्या या कुल योग (कुल योग) को जोड़ने के लिए संख्यात्मक कॉलम के जोड़ की गणना की जाय। यह सुनिश्चित करने के लिए कि डाटा पूर्ण है, पूर्णता नियंत्रण उपाय किये जाने चाहिए उदाहरणतः वैयक्तिक कर दाताओं द्वारा संग्रहीत कर कार्यालय में कुल संग्रहीत कर में जोड़ना चाहिए। लेखापरीक्षक को यह बताते हुए एक प्रमाण पत्र प्राप्त करना चाहिए कि डाटा प्राप्त करते समय लेखापरीक्षित सत्व के आईटी सिस्टम के अनुसार पूर्ण तथा उसी के जैसा है। उक्त प्रमाण-पत्र का संकेतात्मक स्वरूप **अनुबंध 2** में उपलब्ध करया गया है। यह सुनिश्चित किया जाना चाहिए कि केवल प्राधिकृत कार्मिक ही लेखापरीक्षकों को डाटा स्रोतों से डाटा ट्रांसफर करें। अप्राधिकृत रूप से डाटा तक एक्सेस को रोकने के लिए ऐसे डाटा तक एक्सेस केवल उपयुक्त एक्सेस कंट्रोल द्वारा होनी चाहिए।

**किसी सत्व से डाटा जो लेखापरीक्षा क्षेत्राधिकार में नहीं है**

---

<sup>4</sup> विस्तृत डाटा प्रबंधन पॉलिसी, भाग-IV 2, डाटा प्रबंधन प्रोटॉकोल यह सुनिश्चित करे कि डाटा में अग्र लिखित विशेषताएं हो: **विश्वसनीयता**- डाटा अपनी प्रक्रिया से सृजित हो। **समग्रता**- डाटा पूर्ण, सटीक, विश्वसनीय हो। **प्रासंगिकता**- उपयुक्त और निर्दिष्ट उद्देश्य के लिए हो। **उपयोगिता**- आसान रूप से प्राप्त हो। **सुरक्षा**- डाटा सुरक्षित हो और केवल प्राधिकृत पार्टियों की पहुँच में हो।

- 2.14 क्षेत्रीय कार्यालयों की डाटा सेट की आवश्यकता हो सकती है जिसका स्वामित्व उनकी लेखापरीक्षा क्षेत्राधिकार के अंतर्गत लेखापरीक्षा किये जाने वाले सत्वों के साथ नहीं है। क्षेत्रीय कार्यालय तब उस संबंधित क्षेत्रीय कार्यालय की सहायता प्राप्त कर सकते हैं, जिसका ऐसे लेखापरीक्षा किये जाने वाले सत्वों पर लेखापरीक्षा क्षेत्राधिकार हो और संबंधित क्षेत्रीय कार्यालय को अपेक्षित डाटा सेट प्राप्त करने में सभी सहायता प्रदान करनी चाहिए।

### **डाटा का स्वामित्व**

- 2.15 डाटा सेट का स्वामित्व लेखापरीक्षित सत्व/तीसरी पार्टी डाटा स्रोत का है और आईएएंडएडी केवल वैश्वसिक क्षमता में इस डाटा को रखता है। एक बार डाटा स्रोतों से डाटा सेट प्राप्त करने पर, एचओडी के पास डाटा सेट का स्वामित्व होना चाहिए और लेखापरीक्षित सत्व में डाटा मालिक के लिए परिकल्पित डाटा की सुरक्षा और गोपनीयता जैसे नियंत्रणों पर कार्य करना चाहिए। डाटा के मालिक की चिंताओं और निर्देशों, यदि कोई है, को सुनिश्चित करना चाहिए और ध्यान में रखना चाहिए। डाटा स्रोतों द्वारा दिया गया डाटा संदर्भ के लिए सुरक्षित रखा जाना चाहिए और सभी विश्लेषण स्रोत डाटा की केवल प्रतियों में किए जाने चाहिए। लेखापरीक्षित सत्व/तीसरी पार्टी के डाटा की सुरक्षा गोपनीयता और डाटा के उपयोग के संबंध में सभी नियमों, प्रक्रियाओं और समझौतों की अनुपालना समय-समय पर आईएएंडएडी द्वारा निर्दिष्ट डाटा सुरक्षा और संरक्षा के पूर्ण फ्रेमवर्क में लेखापरीक्षा द्वारा सुनिश्चित की जानी चाहिए।

### **डाटा सुरक्षा**

- 2.16 इलेक्ट्रॉनिक रिकॉर्ड के मामले में, कई प्रतियां बनाने के लिए, डाटा संशोधन, डाटा हटाना आदि मैनुअल रिकॉर्ड की तुलना में आसान और तीव्र है। लेखापरीक्षा सत्व पर लागू डाटा सुरक्षा प्रोटोकॉल अपेक्षित डाटा सेटों के प्रहस्तन के लिए लेखापरीक्षकों द्वारा अपनाया जाय। तथापि, डाटा विश्लेषण परिणामों को आईएएंडएडी के निर्दिष्टानुसार देखा जाय।
- 2.17 डाटा प्रहस्तन के समय, आधारभूत दृष्टिकोण डाटा तक एक्सेस वाले कार्मिकों की संख्या और डाटा एक्सेस करने वाले कार्मिकों का देखने को लिए आवश्यकतानुसार सीमित किया जाना चाहिए। डाटा स्रोत मालिक और लेखापरीक्षक के बीच शेयर किये

गये सभी डाटा के पूर्ण और कालक्रमानुसार रिकॉर्ड को परिवर्तन रहित और सुरक्षित रूप से स्टोर किया जाना चाहिए। यह सुनिश्चित किया जाना चाहिए कि डाटा विश्लेषण हेतु प्रयुक्त कम्प्यूटर इंटरनेट से नहीं जुड़े हैं।

- 2.18 लेखापरीक्षित सत्व से प्राप्त डाटा की संवेदनशीलता के कारण, इसे लेखापरीक्षकों से किसी भी प्रकार के अप्राधिकृत प्रकटन से बचने के लिए पूर्ण सावधानी बरतनी चाहिए। सरकार<sup>5</sup> के सूचना सुरक्षा उपाय, जिन्हें आईएएंडएडी की सूचना प्रणाली सुरक्षा हैंडबुक में निर्दिष्ट किया गया है, लेखापरीक्षक और डाटा स्रोत मालिक के बीच एक विशेष समझौते के साथ डाटा की गोपनीयता और सुरक्षा को सुनिश्चित करने के लिए अपनाया जाना चाहिए।

### **डाटा विश्वसनीयता**

- 2.19 डाटा को विश्वसनीय माना जाता है जब प्रस्तुत किए जा रहे पैरामीटर को सटीक रूप से कैप्चर करता है। डाटा विश्वसनीयता, प्रमाणिकता, अखण्डता, प्रासंगिकता और डाटा की प्रयोज्यता का एक कार्य है। डाटा विश्वसनीयता डाटा उत्पादन/कैप्चर की विधियों के कारण प्रभावित हो सकती है। चूंकि आईए एंड एडी को अन्य स्रोतों से उत्पन्न डाटा पर निर्भर होना पड़ता है इसलिए यह महत्वपूर्ण है कि प्रत्येक डाटा स्रोत की विश्वसनीयता को पहले से समझा जाये ताकि इस के उपयोग में पर्याप्त सावधानी रखी जा सके।
- 2.20 सामान्य तौरपर लेखापरीक्षकों के पास लेखापरीक्षण योग्य सत्व से डाटा प्राप्त करने के दौरान डाटा की विश्वसनीयता सुनिश्चित करने के लिए सीमित साधन होते हैं क्योंकि विश्वसनीयता केवल लेखापरीक्षा प्रक्रिया में डाटा के उपयोग के बाद ही निर्धारित की जा सकती है जब विश्लेषण से आंतरिक असंगतता अथवा अपूर्णता का पता चल सकेगा। तथापि, लेखापरीक्षकों को डाटा विश्वसनीयता के बारे में सचेत/सतर्क रहने की जरूरत है और लेखापरीक्षा योग्य सत्वों से डाटा प्राप्त करते समय सावधानी बरतने की आवश्यकता है। सामान्य तौरपर यदि मैनुअल और आईटी प्रणाली समानांतर काम कर रहे हैं, तो डाटा में त्रुटियों की संभावनाएँ अधिक हैं। समानरूप से मैनुअल डाटा प्रविष्टि वाली एक एमआईएस प्रणाली उन प्रणालियों की तुलना में कम विश्वसनीय होती है जहाँ एमआईएस डाटा एक आईटी प्रणाली द्वारा सीधे तौरपर

<sup>5</sup> दिनांक 14 अक्टूबर 2014 को सरकारी नेटवर्क पर आईटी डिवाइस के प्रयोग पर दिशा निर्देश

<http://meity.gov.in/writereaddata/files/Guidelines%20for%20Use%20of%20IT%20Devices%20on%20Government%20Network%200.pdf>



उत्पन्न होता है। आईटी प्रणाली की सूचना प्रणाली लेखापरीक्षा, यदि कोई पहले आयोजित की गई है, तो डाटा विश्वसनीयता पर अंतर्दृष्टि प्रदान कर सकती है।

- 2.21 लेखापरीक्षकों को उन उद्देश्यों के बीच स्पष्ट रूप से अंतर करने की आवश्यकता होती है जिनके लिए डाटा विश्वसनीयता पर विचार करते समय डाटा सेट का उपयोग किया जाएगा। नियोजन के दौरान विस्तृत अंतर्दृष्टि प्राप्त करने के लिए नियोजित डाटा सेट की तुलना में लेखापरीक्षा निष्कर्षों के समर्थन के लिए लेखापरीक्षा प्रमाण के तौरपर प्रयोग के लिए नियोजित डाटा सेट के लिए डाटा विश्वसनीयता पर विचार करना अधिक महत्वपूर्ण होगा। बड़ी डाटा प्रबंधन नीति विभिन्न तृतीय दल डाटा स्रोतों का उल्लेख करती है जिनका प्रयोग आईए एवं एडी में लेखापरीक्षा के लिए हो सकता है। जबकि तृतीय दल डाटा लेखापरीक्षा नियोजन प्रक्रिया को मज़बूत कर सकता है, वहीं लेखापरीक्षक को लेखापरीक्षा प्रमाण के तौरपर ऐसे डाटा स्रोतों का प्रयोग करने के दौरान व्यावसायिक निर्णय का प्रयोग करना चाहिए और यह सुनिश्चित करना चाहिए कि यह भारत के सीएजी के लेखापरीक्षण मानकों के अनुसार निर्धारित मापदंड को पूरा करता है। उदाहरणतः स्वच्छता के संबंध में अकादमिक संस्थान का सर्वेक्षण डाटा को सेक्टर में मामलों का पता लगाने के लिए प्रयुक्त किया जा सकता है और लेखापरीक्षा नियोजन चरण में अन्य पैरामीटर के साथ सैम्पलिंग प्रक्रिया (उच्च जोखिम/निम्न जोखिम प्रशासनिक इकाइयों का पता लगाना) में फीड किया जा सकता है। तथापि, सर्वेक्षण डाटा के विश्लेषणात्मक परिणाम लेखापरीक्षा प्रमाण के तौरपर प्रयुक्त हो सकते हैं जो इस पर निर्भर करता है कि क्या वे आईए एवं एडी के लिए निर्धारित लेखापरीक्षा प्रमाण की शर्तों, मापदंड एवं मानकों को पूरा करते हैं।

### डाटा तैयार करना

- 2.22 अभिज्ञात डाटासेट, जैसे भी उपलब्ध हो, विश्लेषण के लिए हमेशा वांछित रूप, आकार या गुणवत्ता में नहीं हो सकते। इसलिए डाटा उपलब्ध फॉर्मेट से वांछित फॉर्मेट में तैयार करना होगा। बाद के विश्लेषण के लिए, डाटा के वांछित फॉर्मेट पर निर्णय लेने के लिए लेखापरीक्षक के लिए डाटा को समझना एक पूर्वअपेक्षा है।
- 2.23 डाटा तैयार करना विश्लेषण उद्देश्यों के लिए डाटा को व्यवस्थित करने की प्रक्रिया है। इसमें रिस्टोरेशन, डाटा आयात करना, डाटाबेस/तालिका/रिकॉर्ड/फील्ड का चयन, डाटासेट में शामिल होने, डाटासेट जोड़ने, सफाई, एकत्रीकरण और लापता मूल्यों, अमान्य मूल्य, आउटलेयर और परिवर्तन के उपचार जैसी विभिन्न गतिविधियां शामिल

हैं। ये गतिविधियाँ या तो एक दूसरे से अन्तर सम्बन्धित या स्वतंत्र चरणों की एक श्रृंखला हो सकती हैं। डाटा तैयार करना परियोजना<sup>6</sup> विशिष्ट चरण है। यद्यपि विस्तृत कदम महत्वपूर्ण ढंग से भिन्न नहीं हो सकते, इसलिए अन्तर्गस्त उपप्रक्रियाओं या कार्यों का क्रम परियोजना के अनुसार अलग-अलग हो सकता है। इसके अलावा, कुछ कदम/कार्यों को बैक ट्रेक करने या दोहराने की आवश्यकता हो सकती है।

### **डाटा का पुनः स्थापन**

- 2.24 डाटा स्रोत से डाटा आगे विश्लेषण के लिए लेखापरीक्षक के कम्प्यूटर में कॉपी और पुनः स्थापित करना चाहिए। डम्प/बैकअप फार्मेट में डाटा के प्रयोग के दौरान, डाटा तालिकाओं को डाटा रिस्टोरेशन प्रक्रिया द्वारा वास्तविक फार्मेट में लाना आवश्यक होगा।

डाटाबेस बैकअप/डम्प फाइल को रिस्टोर करने से पूर्व कुछ मूलभूत सूचना जैसे कि डाटाबेस सॉफ्टवेयर संस्करण, प्रचालन प्रणाली, डाटाबेस आकार आवश्यक है। इस सूचना के आधार पर, यदि पहले उपस्थित न हो, बैकअप/डम्प फाइल को रिस्टोर करने के लिए एक परिवेश बनाना चाहिए। डाटाबेस रिस्टोरेशन में पर्याप्त डाटाबेस के तकनीकी ज्ञान की आवश्यकता होती है चूंकि एक डाटा को रिस्टोर करने के दौरान उठाए जाने वाले कदम डाटाबेस सॉफ्टवेयर के अनुसार बदल सकते हैं। जबकि निम्न संस्करण बैकअप/डम्प फाइल को डाटाबेस सॉफ्टवेयर के उच्च संस्करण में रिस्टोर करना संभव है, इसमें अनुकूलता मामले हो सकते हैं जिनकी डाटाबेस प्रशासक से पुष्टि करनी चाहिए।

### **रूचि की तालिकाओं/फील्ड की पहचान**

- 2.25 परिकलनीय गति एवं क्षमता को इष्टतम करने के उद्देश्य से यह आवश्यक है कि केवल प्रासंगिक डाटा परिवर्तियों को ही विश्लेषणात्मक उद्देश्यों के लिए रखा जाए। रूचि के प्रासंगिक फील्ड/तालिका/परिवर्ती का पता लगाना अधिकतम सावधानी से करना होगा चूंकि यदि बाद के चरण में कोई भी अतिरिक्त फील्ड/तालिका/परिवर्ती को प्रासंगिक पाया गया तो सभी प्रक्रियात्मक चरणों को फिर से दोहराया जा सकता है।

---

<sup>6</sup> यहाँ परियोजना, या तो एक लेखापरीक्षा करने के दौरान या डाटा स्रोतों से प्राप्त डाटा का विश्लेषण करने के दौरान एक डाटा विश्लेषण परियोजना है, जो कि आवश्यक रूप से लेखापरीक्षा से जुड़ी नहीं है।

## विश्लेषणात्मक उपकरण में प्राप्त करना

- 2.26 अधिकतम विश्लेषणात्मक टूल्स के पास फ्लैट फाइल को सॉफ्टवेयर में पढ़े जाने या डाटाबेस से जुड़ने और तालिकाओं को पढ़ने का विकल्प है। कुछ विश्लेषण सॉफ्टवेयर केवल कॉलम/तालिकाएँ को प्राप्त करने और प्लेटफॉर्म में फाइल के पढ़े जाने के पूर्व डाटा टाइप बदलने का विकल्प प्रदान करते हैं। स्वयं विश्लेषण टूल डाटा को साफ एवं वर्धन के विभिन्न विकल्प प्रस्तावित करते हैं। डाटा की गुणवत्ता एवं मात्रा पर निर्भर करते हुए, लेखापरीक्षक एक स्प्रेडशीट या आरडीबीएमएस में विश्लेषण प्लेटफॉर्म या बाहर के तहत डाटा क्लीनिंग/वर्धन का चयन कर सकता है। प्राप्त करने और डाटा की सफाई के चरण डाटासेट और उपयुक्त टूल की उपलब्धता पर निर्भर करते हुए पहले या एक दूसरे के बाद में आ सकते हैं।

## डाटा फाइलों की मर्जिंग एवं स्पिलटिंग

- 2.27 डाटा स्रोतों से प्राप्त डाटा विभिन्न अवधियों या स्थानों से संबंधित हो सकता है अथवा साधारण तौरपर विभिन्न भागों में बाँटा जा सकता है। डाटा को विश्लेषण के अधीन लाने के लिए, डाटासेटों को एक में मर्ज करना आवश्यक होगा। यह डाटा फाइलों को जोड़ने द्वारा किया जा सकता है। समानरूप से, एक सत्त्व से संबंधित डाटा सेट में विभिन्न कार्यों/पैरामीटर के विवरण होते हैं। ऐसे मामलों में, सभी पैरामीटरों को विश्लेषण के लिए एक फाइल में लाने के लिए सभी डाटा फाइलों को जोड़ा जाय।
- 2.28 डाटा फाइलें, डाटा फाइलों को डाटा सेट कम करने के लिए भी विभाजित किया जा सकता है, जिससे कार्यकुशल विश्लेषण में सहायता मिलती हैं। फाइलों को रिकॉर्ड की संख्या या पैरामीटर की संख्या के आधार पर विभाजित किया जाय। फाइलों की मर्जिंग और विभाजन आरडीबीएमएस या डाटा विश्लेषण संबंधी उपकरण द्वारा किया जा सकता है।

## डाटा क्लीनिंग

- 2.29 अच्छी गुणवत्ता का डाटा जो कि साफ, पूर्ण और त्रुटि रहित हो अच्छे विश्लेषण के लिए आवश्यक है। डाटा अपमार्जन, डाटा की सफाई या डाटा सक्रबिंग एक रिकॉर्ड सेट, तालिका या डाटाबेस से भ्रष्ट या अशुद्ध रिकार्ड का पता लगाने और सही करने या हटाने की प्रक्रिया होती है। यह अपूर्ण गलत, अशुद्ध या डाटा के गैर प्रासंगिक भाग

का पता लगाने और अशुद्ध या भ्रष्ट डाटा का प्रतिस्थापन, संशोधन या फिल्टर करने से संबंधित है। डाटा क्लीनिंग की प्रक्रिया में मुद्रण की गलती हटाने या सत्त्वों की ज्ञात सूची के प्रति वैल्यू को प्रमाणित करना और सही करना या एक वैध डाटा सेट की जांच शामिल हो सकता है। डाटा क्लीनिंग में रिकार्डों को अस्वीकार करना या सही करना और किसी भी अवैध वैल्यू की मौजूदगी का सत्यापन करना शामिल है।

- 2.30 डाटा संवर्धन भी एक डाटा अपमार्जन प्रक्रिया है जहाँ डाटा को संबंधित सूचना जोड़ने के द्वारा अधिक पूर्ण बनाया जाता है। इसमें डाटा का समानीकरण और मानकीकरण शामिल है। उदाहरणतः एक बैंक का नाम बैंक कोड से जोड़ना डाटा की गुणवत्ता में संवर्धन करता है। समानरूप से, छोटे कोड (एसटी, आरटी आदि) का वास्तविक शब्दों (गली, सड़क आदि) में समानीकरण किया जा सकता है। डाटा का मानकीकरण एक संदर्भ डाटा सेट को एक नये मानक उदाहरणतः मानक कोड का उपयोग में परिवर्तित करने का साधन है।

#### **गायब मूल्य और अन्य डाटा बनाने के चरण**

- 2.31 मूल्य गायब तब होता है जब डाटासेट में एक फील्ड में परिवर्ती के लिए कोई डाटा मूल्य उपलब्ध न हो। यह एक सामान्य घटना है, जोकि डाटासेट की प्रतिनिधिकता को कम करती है और डाटा से निकाले गए अनुमान एवं निष्कर्षों को विकृत कर सकती है। गायब मूल्य यादृच्छिक रूप से या एक पैटर्न में हो सकते हैं। शेष डाटा को उचित ढंग से संभालने के लिए उनके कारणों और गायब मूल्य की प्रवृत्ति की समझ महत्वपूर्ण है। गायब मूल्य की प्रकृति पर आधारित, डाटा सेट को या तो गायब मूल्य को काटे जाने या उन्हें अन्य मूल्यों जैसे कि उपलब्ध मूल्यों का माध्य, औसत या मोड के साथ सौंपे जाने द्वारा संभालना चाहिए।
- 2.32 अन्य डाटा तैयार करने के चरणों में अनचाहे कॉलम को हटाने, फार्मेटिंग और विभिन्न कॉलम का पुनः नामकरण एवं अतिरिक्त कॉलम डालना (जैसे कि एक अतिरिक्त वर्ष कॉलम चलन विश्लेषण के लिए डालना) शामिल है।

#### **डाटा एकीकरण: बहुल डाटाबेस को लिंक करना**

- 2.33 डाटा एकीकरण एक प्रक्रिया है जहाँ पर उसी डाटा स्रोत के तहत विभिन्न डाटा स्रोतों या विभिन्न तालिकाओं से संग्रहीत डाटा को विश्लेषण के लिए अंतिम डाटासेट प्राप्त करने के लिए जोड़ा जाता है। किसी भी सामान्य फील्ड जैसे कि यूनीक कस्टमर

आईडी, बिल संख्या या गाँव का नाम आदि पर आधारित डाटा के विभिन्न स्रोतों को एकीकृत किया जा सकता है। उदाहरणतः यह समझने के लिए कि क्या एक निश्चित सामाजिक सुरक्षा योजना के तहत लाभार्थियों का कवरेज जनसंख्या वितरण से सहसंबंधित है, लाभार्थी डाटा को ज़िला स्तर तालुका स्तर, या आगे के कणमय स्तरों पर संगणना स्तर से जोड़ा जाय (जोड़ा गया है)। विभिन्न डाटा स्रोतों के मेटा डाटा<sup>7</sup> की समझ डाटा एकीकरण की प्रक्रिया में सहायता करेगी।

- 2.34 बहुल डाटा सेट को लिंक करने के दौरान, डाटासेट में एक सामान्य फ़िल्ड होना आवश्यक नहीं है चूंकि तुलनाओं को सक्षम करने के लिए उच्च स्तर पर डाटा को जोड़ा जा सकता है। उदाहरणतः एक व्यक्तिगत लाभार्थी को पेंशन लाभार्थी डाटाबेस और बीपीएल डाटाबेस से जोड़ना संभव नहीं होगा, डाटा को गाँव/ब्लॉक/ज़िला स्तर पर गाँव का पता लगाने के लिए जोड़ा जा सकता है जहाँ पर इन संख्याओं के बीच मेल न हो। ऐसे बेमेलपन के कारण का तब लेखापरीक्षा द्वारा मूलभूत जाँच के दौरान अन्वेषण किया जा सकता है।

### 3. डाटा विश्लेषण एवं मॉडलिंग

#### डाटा विश्लेषण दृष्टिकोण

- 3.1 विभिन्न विश्लेषण दृष्टिकोणों का उपयोग करते हुए अंतर्दृष्टि प्राप्त करने के लिए उचित तैयारी के बाद डाटा का विश्लेषण किया जाता है। डाटा विश्लेषण में निम्नलिखित तरीकों का उपयोग किया जा सकता है:

- **वर्णनात्मक विश्लेषण** “क्या हुआ है” का उत्तर देने की कोशिश करता है। वर्णनात्मक विश्लेषण में, कच्चे डाटा का संक्षेपण किया जाता है ताकि उपयोगकर्ता द्वारा इसे समझा जा सके। वर्णनात्मक विश्लेषण संगठन में हुए पिछले लेनदेन की समझ प्रदान करता है। वर्णनात्मक विश्लेषण में व्यक्तिगत लेनदेन का एकत्रीकरण शामिल है और इस प्रकार एक बड़े परिप्रेक्ष्य में व्यक्तिगत लेनदेन के लिए अर्थ और संदर्भ प्रदान करता है। इसमें संख्यात्मक या दृश्य विवरणों के माध्यम से डाटा का संक्षिप्तीकरण शामिल है।

<sup>7</sup> मेटा डाटा अन्य डाटा सेट का डाटा है। इसमें डाटासेट पर सूचना इस प्रकार निहित है जो डाटा सेट का पता लगाने को आसान बनाती है।

- **डायग्नोस्टिक विश्लेषण** वर्णनात्मक विश्लेषण का एक उन्नत रूप है “ऐसा क्यों हुआ” प्रश्न का उत्तर देने की कोशिश करता है। डायग्नोस्टिक विश्लेषण में उनके व्यवहार और अंतर्निहित कारणों के साथ सापेक्ष डाटा सेट और विशिष्ट लेनेदेन/लेनदेन सेट की पहचान के बीच संबंधों को समझना शामिल है। विभिन्न घटनाओं के कारणों को समझने के लिए इस प्रयास में सहसंबंध और सांख्यिकीय तकनीक की सहायता करती है।
- **भविष्यवाणीत्मक विश्लेषण**, जैसा कि नाम से अर्थ है, पूर्व आंकड़ों के आधार पर, भविष्यवाणी करने की कोशिश करता है, “क्या होगा”, “यह कब होगा”, “यह कहां होगा” एक निश्चित सीमा तक, एक गतिविधि के भविष्य के परिणामों का पता लगाने के लिए विभिन्न पूर्वानुमान और आंकलन तकनीकों<sup>8</sup> का उपयोग किया जा सकता है।
- **निर्देशात्मक विश्लेषण** भविष्यवाणीत्मक विश्लेषण से अधिकार में लेता है और लेखापरीक्षक को इनपुट के तौरपर संभावित कार्यों की सीमा को ‘निर्धारित’ करना अनुमत करता है जैसे कि भविष्य में आऊटपुट को वांछित समाधान में बदला जा सकता है। निर्देशात्मक विश्लेषण में विभिन्न इनपुट व्यवधानों के आधार पर बहुल भविष्य परिदृश्यों का पता लगाया जा सकता है।

### डाटा विश्लेषण तकनीकें

3.2 उपरोक्त दृष्टिकोणों का लाभ उठाने के लिए डाटा विश्लेषणात्मक तकनीकों का उपयोग किया जाता है। विश्लेषणात्मक तकनीकें जो वर्णनात्मक और डायग्नोस्टिक दृष्टिकोण का उपयोग करती हैं, लेखापरीक्षक को लेखापरीक्षा करने योग्य सत्व को समझने और उसमें समस्याओं की पहचान करने में मदद करेगी। पुनरावृत्ति की तरह एक भविष्यसूचक तकनीक परिवर्तियों के दूसरे सेट में परिवर्तन के आधार पर एक (या अधिक) परिवर्ती के व्यवहार को समझने में मदद करेगी। इन विश्लेषणात्मक

<sup>8</sup> पूर्वानुमान और अनुमानित तकनीकों में, पिछले डाटा उपलब्ध ज्ञान दस्तावेजों मान्यताओं और पता लगाए गए जोखिम का उपयोग शामिल हैं और संचालन अनुसंधान और मात्रात्मक तकनीक विषयोंका भाग हैं।

तकनीकों को मोटे तौर पर सांख्यिकीय और विजुअली<sup>9</sup> के रूप में वर्गीकृत किया जा सकता है।

- **सांख्यिकीय तकनीक** डाटासेट के विषय में इनसाइट प्राप्त करने के लिए सांख्यिकीय उपायों का उपयोग है।
- **दृश्यकरण तकनीक डाटासेट** की समझ तथा उनकी इनसाइट प्राप्त करने के लिए दृश्यों, ग्राफो तथा चार्टों का उपयोग है।

विभिन्न सांख्यिकीय तथा दृश्य तकनीकों के संयोजन को सामान्य रूप से डाटा विश्लेषण के लिए उपयोग किया जाता है।

- 3.3 यह डाटा विश्लेषण सॉफ्टवेयर में उक्त वर्णित दृष्टिकोणों के लिए अन्तर्निहित कलन विधि है। यद्यपि, डाटा विश्लेषण तकनीकों के अनुप्रयोग के लिए कोई वर्णित तर्कसंगत कार्रवाई नहीं है जिसका जूम आउट-जूम इन-फिल्टर एप्रोच के रूप में व्यापक रूप से वर्णन किया जा सकता है। डाटा को पहले गहन स्तर पर डाटा की समझ लेने के लिए एक ड्रिल डाउन द्वारा अनुसरित विहंगम दृष्टि से समझा जाता है। इसके पश्चात्, एक फिल्टर किया जाता है अथवा यदि आवश्यक हो तो परिणामों अथवा अपवादों के लिए एक क्वेरी की जाती है। उदाहरणार्थ, सम्पत्ति कर मांग तथा संग्रहण वाले एक डाटा सेट में, कोई भी कर मांग/संग्रहणों की औसत रेंज तथा करदाताओं की रेंज तथा जूम आउट के साथ विभिन्न जोनों में उनके वितरण को समझा जा सकता है। जूम इन के साथ कोई भी विभिन्न जोनों में कर भुगतानों के अस्थाई तथा प्रभेद पैटर्नों के बीच सह-संबंध को समझ सकता है। इसके पश्चात्, कर बकाया को निर्धारित अधिक जोखिम क्षेत्र फिल्टर किया जा सकता है। इसके अलावा, यह देखने के लिए एक प्रतिगमन विश्लेषण भी किया जा सकता है कि भविष्य में किस जोन के पास अधिकतम बकाया होंगे।

### **सांख्यिकीय तकनीकें**

- 3.4 एक बार डाटा तैयार हो जाए तो प्रथम स्टेप के रूप में डाटा सेट में वर्णित प्रत्येक सांख्यिकीय माप के साथ किसी तरीके से डाटा को संक्षिप्त करने के लिए डाटासेट की

---

<sup>9</sup> यद्यपि अधिकांश ग्राफिक और चार्ट जैसे विजुअलाइज़ेशन तकनीक अनिवार्य रूप से सांख्यिकीय हैं ये इन मायने में अलग हैं कि समझ केवल सांख्यिकीय उपायों से नहीं बल्कि तुलना विश्लेषण और विजुअली अंतर्दृष्टि उत्पन्न करने से प्राप्त होती है।

व्याख्यात्मक सांख्यिकीय प्रस्तुत की जा सकती है। इसे लाइन ग्राफ, हिस्टोग्राम अथवा स्कैटर डायग्राम जैसी सामान्य ग्राफिक प्रस्तुति द्वारा किया जा सकता है। उदाहरणार्थ, केन्द्रीय प्रवृत्ति का मापन एक विशिष्ट मानदण्ड अथवा वेरिएबल के संदर्भ में एक सत्त्व तथा इसके सिद्धांतों के अपेक्षित सामान्य व्यवहार को वर्णित करते हैं। स्प्रेड के उपाय डाटा प्वाइंटो के वितरण को दर्शाते हैं। दो अथवा अधिक परिवर्तों के बीच संबंध को सहसंबंध<sup>10</sup> तथा प्रतिगमन<sup>11</sup> की तकनीकों का उपयोग करके उजागर तथा स्थापित किया जा सकता है। महत्वपूर्ण मानदण्डों के निर्धारण अथवा पृथक्करण को प्रतिगमन, घटक विश्लेषण<sup>12</sup> अथवा कारक विश्लेषण<sup>13</sup> का उपयोग करके किया जा सकता है। क्लस्टरिंग<sup>14</sup> तथा क्लासिफिकेशन<sup>15</sup> का उपयोग एक अथवा अधिक समानता पर आधारित डाटासेट में ग्रुप को निर्धारित करने के लिए किया जा सकता है। विभिन्न सांख्यिकीय जांचो से परिणाम को डाटासेट की अंतिम समझ प्राप्त करने के लिए एक साथ पढ़ना आवश्यक है।

### डाटा दृश्यकरण

3.5 डाटा दृश्यकरण निम्नलिखित दो भिन्न प्रयोजनों के लिए काम आता है:

- **अन्वेषी डाटा विश्लेषण (ईडीए):** यह प्रायः दृश्यकरण प्रक्रिया से डाटा सेटों का उनकी प्रमुख विशेषताओं को संक्षिप्त रूप देने के लिए विश्लेषण करने का एक दृष्टिकोण है। प्राथमिक रूप से, ईडीए यह देखने के लिए किया जाता है कि डाटा हमें सांख्यिकीय विश्लेषण तथा मॉडलिंग से अधिक क्या बता सकता है।
- **निष्कर्ष/रिपोर्टिंग की सूचना:** डाटा से प्राप्त इनसाइट को उच्च प्रबंधन अथवा लेखापरीक्षा रिपोर्टों के पाठक जैसे उपयोगकर्ताओं को सूचित किया जा सकता है।

<sup>10</sup> दो वेरिएबल और  $-1$  से  $+1$  के बीच रेंजों के मध्य संपर्क की मजबूती को मापने के लिए सहसंबंध का प्रयोग किया जाता है।

<sup>11</sup> प्रतिगमन विश्लेषण एक अंकीय स्पष्टीकरण देता है कि वेरिएबल को जोड़ने में स्वतंत्र वेरिएबल में दिए गए निर्भर वेरिएबल भविष्यवाणी को कैसे सक्षम बनाते हैं।

<sup>12</sup> प्रमुख घटक विश्लेषण उस एक छोटे सेट के लिए अन्तर-संबंधित वेरिएबल की संख्या को कम करने के लिए लक्षित है जो सम्पूर्ण अस्थिरता को वर्णित करता है।

<sup>13</sup> कारक विश्लेषण उन वेरिएबल को एक साथ ग्रुप करने तथा संक्षिप्त करने के लिए लक्षित है जो सहसंबंधित हैं जिससे डाटा कमी सक्षम होती है।

<sup>14</sup> क्लस्टर विश्लेषण सामान्य विशेषताओं पर आधारित ग्रुप व्यक्तिगत/वेरिएबलस के लिए उपयुक्त एक मल्टीवेरिएट तकनीक है ([www.statstutor.ac.uk](http://www.statstutor.ac.uk) देखें)

<sup>15</sup> डाटा में प्रस्तुत कुछ सामान्य विशेषताओं के अनुसार एक समान ग्रुप अथवा क्लास के अन्दर डाटा को व्यवस्थित करने की प्रक्रिया वर्गीकरण कहलाती है।

(Ref:<http://www.emathzone.com/tutorials/basic-statistics/classification-of-data.html#ixzz4r2Rlugdu>)



डाटा दृश्यकरण डाटा विश्लेषण परिज्ञान को सूचित करने के लिए एक सशक्त तकनीक है।

3.6 डाटा दृश्यकरण निम्नलिखित एक या अधिक उद्देश्यों को प्राप्त करने के लिए लक्षित है:

- बोधगम्यता: सूचना तथा सह-संबंधता को सहज रूप से समझ योग्य बनाती है।
- व्यापकता: चयनित रिपोर्टिंग के प्रति सम्पूर्ण चयनित डाटा सेट/नमूना आकार के लिए विशेषता/सूचना प्रस्तुत करती है।
- संकेद्रित संप्रेषण: संक्षिप्त तथा 'टू दी प्वाइंट' संप्रेषण को सुविधाजनक बनाता है।
- घटती जटिलता: डाटा की अधिकतम संख्या की प्रस्तुति को सुगम बनाती है।
- स्थापित पैटर्न तथा सहसंबंध: डाटा में पैटर्न तथा संबंध के निर्धारण को सक्षम बनाता है।
- विश्लेषण: 'कार्यप्रणाली' के बजाय 'तात्पर्य' पर विचार को बढ़ावा देता है। यह संप्रेषण के लिए प्रक्रिया की बजाय सम्प्रेक्षित किए जा रहे निष्कर्ष की आवश्यकता पर ध्यान केन्द्रित करता है।

डाटा दृश्यकरण के सिद्धांतों के लिए डाटा दृश्यकरण तथा इन्फोग्राफिक के उपयोग हेतु आईए एंड एडी व्यवसायी गाइड को संदर्भित किया जाना चाहिए।

3.7 यह नोट किया जाना चाहिए कि एक एकल तकनीक डाटा की व्यापक समझ नहीं देगी। एक लेखापरीक्षक को इनसाइट प्राप्त करने के लिए सांख्यिकीय तथा चित्रण तकनीक के संयोजन का उपयोग करना चाहिए। तकनीकों की उपयुक्तता डाटासेट तथा लेखापरीक्षक के उद्देश्य पर निर्भर करती है।

### **नमूना सांख्यिकीय के बजाय पोपुलेशन सांख्यिकीय**

3.8 आधुनिक डाटा विश्लेषण उपकरणों की सहायता से, एक सम्पूर्ण डाटा सेट का विश्लेषण करना संभव है। इस प्रकार, पोपुलेशन (डाटा सेट में सम्मिलित सभी सव्यंवहार) पर निष्कर्ष नमूनों के माध्यम से निष्कर्ष निकालने की बजाय डाटा सेट के सभी सव्यंवहारों का विश्लेषण करके दिया जा सकता है। तथापि, यदि डाटा सेट आईटी सिस्टम द्वारा ग्रहित पूर्ण व्यवसाय प्रक्रिया का पूर्ण रूप से प्रतिनिधि नहीं है तो वास्तविक जांच की आवश्यकता होगी।

## डाटा विश्लेषण उपकरण

- 3.9 डाटा विश्लेषण एक बहु स्तरीय प्रक्रिया है जिसमें प्रत्येक स्तर पर विभिन्न आवश्यकताओं के साथ मॉडल की तैयारी, विश्लेषण तथा निर्माण सम्मिलित है। इस प्रयोजन हेतु कई सशक्त ओपन स्रोत<sup>16</sup> तथा स्वामित्व साफ्टवेयर<sup>17</sup> उपलब्ध हैं। किसी एकल उपकरण को सभी विश्लेषण अथवा डाटा निष्कर्षण आवश्यकताओं के लिए व्यापक अथवा उपयुक्त नहीं कहा जा सकता। डाटा प्रस्तुति में कुछ उपकरण उपयोगी है जबकि उनमें डाटा दृश्यकरण का अभाव पाया जा सकता है। इसी प्रकार, इसमें सशक्त दृश्यकरण विशेषताओं के साथ ऐसे उपकरण है जिनमें अग्रिम सांख्यिकीय विश्लेषण करने की सक्षमता का अभाव है।
- 3.10 चूंकि लेखापरीक्षक किसी ओपन स्रोत अथवा स्वामित्व साफ्टवेयर का पता लगा सकता है तथा उसे अपना सकता है अतः उपकरण तथा डाटा सुरक्षा बनाए रखने के संदर्भ में उचित देखभाल की जानी चाहिए। एक नए विश्लेषण उपकरण को अपनाते समय, एचओडी को वित्तीय तथा मानव स्रोतों के अनुसार उपकरण के स्थायित्व के मामले पर विचार करना चाहिए। उपकरण की मापनीयता (डाटा सेटों का आकार तथा विविधता की तुलना में) को भविष्य में उपकरण की उपलब्धता पर विचार किए बिना ध्यान में रखने की आवश्यकता है। एचओडी को यह भी सुनिश्चित करना होगा कि लेखापरीक्षित सत्वों के डाटा सेट अथवा कोई अन्य संवेदनशील डाटासेट को अनधिकृत व्यक्तियों/सत्वों के साथ डाटा विश्लेषण साफ्टवेयर के सर्वर/क्लाउड पर्यावरण में शेयर नहीं किया जाता। पर्याप्त सतर्कता के तरीके से जब भी नए उपकरण के उपयोग को एक कार्यालय में औपचारिक रूप दिया जाए तो उसके लिए सीडीएमए से स्वीकृति प्राप्त की जाए।

## डाटा विश्लेषण परिणाम

3.11 डाटा विश्लेषण के परिणाम निम्नलिखित रूप में हो सकते हैं:

- लेखापरीक्षा की सभी जानकारी
- लेखापरीक्षा साक्ष्य

---

<sup>16</sup> नाइम ([www.knime.org](http://www.knime.org)), आर ([www.r-project.org](http://www.r-project.org)), पाइथॉन ([www.python.org](http://www.python.org)), वेका, रेपिडमाइनर, एसपीएजीओ कुछ ओपन स्रोत उपकरण हैं।

<sup>17</sup> एसएसएस, टेबलेयू, एमएस पावर बीआई 17, टिडको, इन्फोमेटिका, आईबीएम विश्लेषणात्मकता, एसपीएसएस, डी3जे, क्लिक आदि कुछ स्वामित्व उपकरण हैं।

## लेखापरीक्षा की सभी जानकारी

- 3.12 लेखापरीक्षक डाटा से सभी जानकारी प्राप्त करने के लिए विभिन्न सांख्यिकीय तथा दृश्यकरण पद्धतियों का प्रयोग करता है। इनसाइट प्राप्त करने के लिए कार्रवाई की कोई वर्णित श्रृंखला नहीं है तथा इसमें पश्च अनुमार्गण तथा पुनरावृत्तियों द्वारा कार्रवाइयों को दोहराकर सम्मिलित किया जा सकता है। लेखापरीक्षक को उन सभी इनसाइट की सूची बनानी चाहिए जिन्हें इस प्रकार प्राप्त किया गया है। यह ध्यान में रखना चाहिए कि जबकि सभी सांख्यिकीय निष्कर्ष कुछ पैटर्न वर्णित करेंगे, वे सभी नए इनसाइट का कारण नहीं बन सकते। डाटा विश्लेषण से प्राप्त इनसाइट डाटा/सत्व की पिछली समझ को सुनिश्चित कर सकते हैं। इसलिए डाटा विश्लेषण से उजागर इनसाइट की सामूहिक रूप से सराहना की जाएगी।
- 3.13 डाटा विश्लेषण प्रक्रिया से प्राप्त परिणामों की सराहना करने के लिए कार्यक्षेत्र की जानकारी आवश्यक है। विभिन्न विश्लेषण तकनीकों का उपयोग करके सृजित निष्कर्ष<sup>18</sup> को सूचीबद्ध तथा उनके मूल्य और महत्वपूर्णता को समझने के लिए कार्यक्षेत्र विशेषज्ञों के साथ उनकी जांच की जानी चाहिए। तब इन इनसाइटों को लेखापरीक्षा के लिए जोखिम क्षेत्रों/ हित क्षेत्रों की पहचान करने के लिए उपयोग किया जाता है। सांख्यिकीय निष्कर्ष तथा निरीक्षण को सूचीबद्ध करने तथा इसका दस्तावेजीकरण करने के लिए एक आदर्श **अनुबन्ध 3** में दिया गया है।

## लेखापरीक्षा साक्ष्य

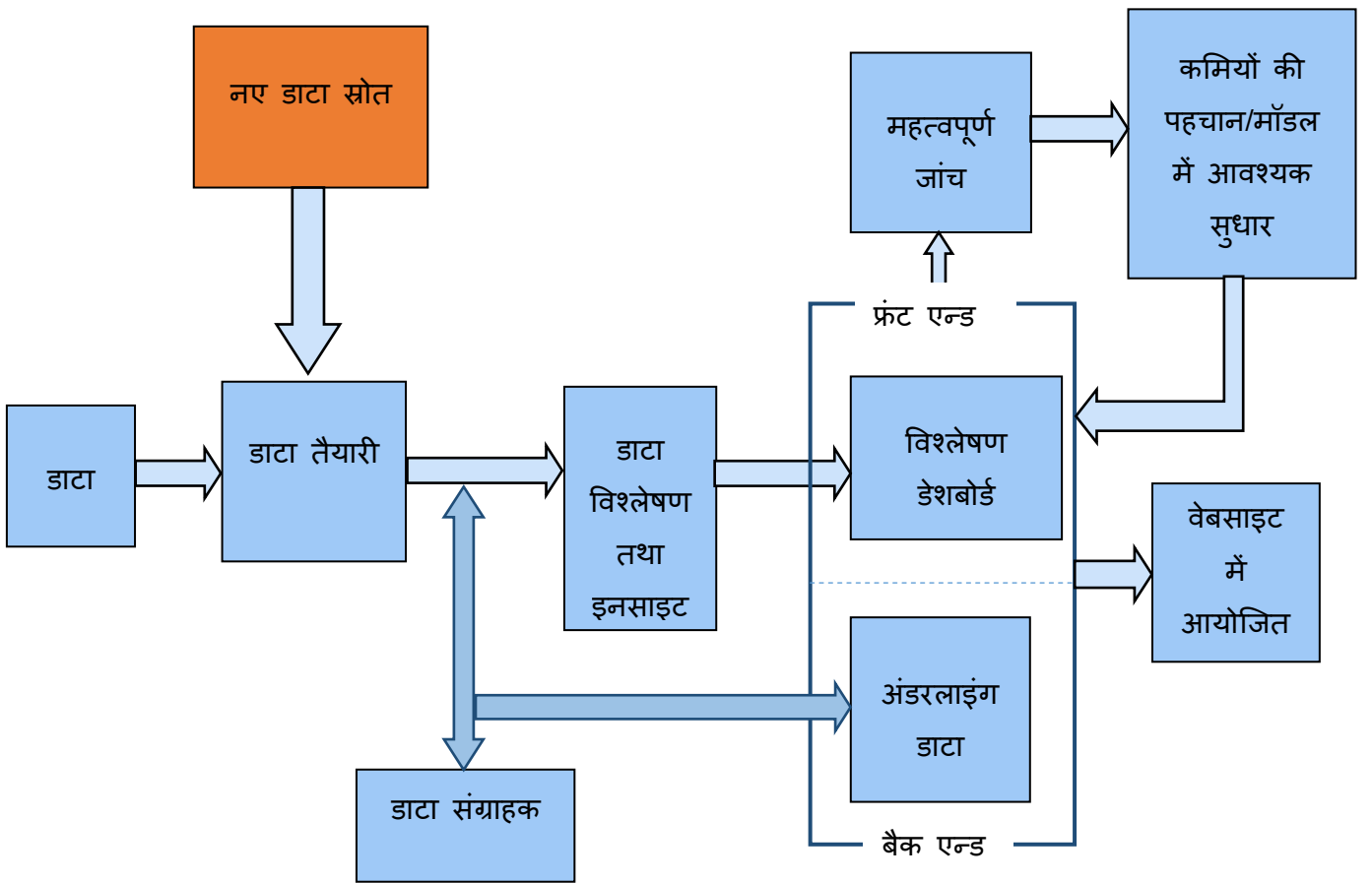
- 3.14 लेखापरीक्षक लेखापरीक्षा निष्कर्षों तथा परिणामों का समर्थन करने के लिए लेखापरीक्षा साक्ष्य के रूप में उनका उपयोग करने के लिए डाटा विश्लेषण परिणामों का मूल्यांकन करने में कुशल निर्णय का प्रयोग करता है। डाटा विश्लेषण परिणामों की महत्वपूर्ण जांचों के माध्यम से एकत्रित साक्ष्य के अन्य रूपों द्वारा पुष्टि की जानी है। डाटा विश्लेषण परिणाम लेखापरीक्षा साक्ष्य के रूप में तब योग्य होंगे जब वे लेखापरीक्षण मानक में वर्णित आवश्यकताओं को पूरा करें।

## डाटा विश्लेषण मॉडल

---

<sup>18</sup> यहां विश्लेषण से निष्कर्ष लेखापरीक्षा निष्कर्ष से भिन्न है। विश्लेषणात्मक निष्कर्षों को विश्लेषण के माध्यम से खोजा जाता है जिससे निरीक्षण होता है। निरीक्षण को एक लेखापरीक्षा निष्कर्ष की पुष्टि करने के लिए महत्वपूर्ण जांच के माध्यम से लेखापरीक्षा में जारी रखा जाता है।

3.15 डाटा विश्लेषण मॉडल विश्लेषण जांच के सेट से संबंधित है जिससे विश्लेषण परिणाम हुए जिन्हें डाटा को अद्यतित/परिवर्तित करके बार-बार उपयोग किया जा सकता है। एक मॉडल का निर्माण यह सुनिश्चित करेगा कि एक बार आवधिक डाटा प्राप्त करने के लिए तन्त्र स्थापित हो जाने पर विशिष्ट डाटासेट पर एक बार किए गए जोखिम विश्लेषण को आगामी वर्षों/अवधियों के लिए एकसमान डाटासेट का उपयोग करके बार-बार प्रयोग किया जा सकता है। डाटा विश्लेषण मॉडल के सृजन की प्रक्रिया को निम्नलिखित फ्लो डायग्राम से स्पष्ट किया गया है:



चित्र 5 - एक डाटा विश्लेषण मॉडल का प्रक्रिया प्रवाह

3.16 एक डाटा विश्लेषण मॉडल बनाने के लिए निम्नलिखित उपायो का पालन किया जाता है:

- विभिन्न स्रोतों से उपलब्ध डाटा को डाटा सेट के रिस्टोरेशन तथा क्लिनिंग के माध्यम से डाटा विश्लेषण के लिए तैयार किया जाता है।

- तैयारी स्तर के बाद, डाटा को डाटा संग्रह में स्टोर किया जाता है।
- इनसाइट को डाटा संग्राहक तथा/अथवा डाटा तैयारी के बाद उपलब्ध डाटा से प्राप्त डाटा सेटो पर विभिन्न विश्लेषण तकनीको का उपयोग करके प्राप्त किया जाता है।
- सुसंगत इनसाइट को डाटा विश्लेषण मॉडल के अन्दर रूपांतरित किया जाता है। मॉडल में इक्वेशन, क्वेरिज, वर्कफ्लो अथवा डेशबोर्ड<sup>19</sup> निहित हो सकते हैं।
- एक बार एक मॉडल तैयार हो जाए तो इसे समीक्षा तथा अनुमोदन के लिए सीडीएमए को प्रस्तुत किया जाना चाहिए।

3.17 डाटा मॉडल को केन्द्रीकृत अथवा विकेन्द्रीकृत डाटा स्रोतो पर विकसित किया जा सकता है:

- **केन्द्रीकृत डाटा स्रोत:** यदि लेखापरीक्षा योग्य सत्व/स्रोत का डाटा केन्द्रीकृत किया जाता है अर्थात यह एक केन्द्रीकृत डाटाबेस के माध्यम से उपलब्ध हो तो पुनः स्टोर किए गए डाटाबेस पर प्रत्यक्ष रूप से एक मॉडल बनाया जा सकता है। वैकल्पिक रूप से, लेखापरीक्षित सत्व के डाटाबेस से सुसंगत तालिकाओं/क्षेत्रों को मॉडल का विकास करने के लिए प्राप्त किया जा सकता है। जब आगामी अवधि के लिए नया डाटा उपलब्ध हो तो इसे मॉडल के अन्दर अपेंडिंग/लोडिंग से पूर्व पुनः स्टोर करने तथा क्लिनिंग (डाटा तैयारी) चरणो के माध्यम से पास करके सम्मिलित किया जाएं।
- **विकेन्द्रीकृत डाटा स्रोत:** यदि लेखापरीक्षित सत्व/स्रोतो के डाटा को विकेन्द्रीकृत किया जाता है (अर्थात प्रत्येक लेखापरीक्षित सत्व से डाटा उप यूनिट एक भिन्न क्षेत्र है जो निर्बाध रूप से जुड़े हुए नहीं है) तो मॉडल को मॉडल के अन्दर अपने डाटा के साथ पूर्व उप-यूनिट के बैकएण्ड डाटा को प्रतिस्थापित करके एक उप यूनिट पर उपयोग किया जाय।
- यदि डाटा को रियल टाइम आधार पर प्राप्त किया जा रहा है तो मॉडल को भी रियल टाइम आधार पर अद्यतित किया जाएगा जिससे निरन्तर लेखापरीक्षण की संभावना होगी।

<sup>19</sup> डेशबोर्ड को एक इंटरफेस (सामान्य रूप से इन्टरक्टिव) के रूप में समझा जा सकता है जिसे अधिक प्रस्तुति योग्य तरीके में शोकेस सूचना/डाटा के लिए उपयोग किया जाता है। डाटा विश्लेषण में, मल्टीपल इनसाइट (ग्राफो/चार्टो के रूप में) को दर्शक की समझ के लिए एक डेशबोर्ड में एक साथ लाया जा सकता है।

- 3.18 एक प्राथमिक मॉडल डाटा विश्लेषण द्वारा दी गई विभिन्न इनसाइट को सम्मिलित करेगा। तथापि, यह संभावना है कि सभी कारको पर विचार न किया गया हो अथवा सभी डाटा विश्लेषण के लिए उपलब्ध नहीं थे; जबकि मॉडल को विकसित किया गया था तथा अन्य इनसाइट तब प्राप्त की जा सकती है जब मॉडल को परिणियोजित किया जाए। डाटा मॉडल को उन अतिरिक्त इनसाइटो के साथ-साथ अधिक उपयुक्त डाटा सेटो के साथ अद्यतित किए जाने की आवश्यकता है जो बाद में उपलब्ध हुआ हो।
- 3.19 इस मॉडल की एक महत्वपूर्ण विशेषता इसकी पुनः उपयोगिता है। एक बार मॉडल के सृजन के बाद इसके डाटा को अद्यतित करके बार-बार उपयोग किया जा सकता है। अतः मॉडल की उपयोगिता डाटा के आवधिक अद्यतन पर निर्भर करेगी। इस प्रकार, डाटा संग्रहण प्रक्रिया को एकबारगी कार्य मानने की बजाय डाटा को वार्षिक रूप से/आवधिक रूप से प्राप्त करने के लिए एक तंत्र बनाए जाने की आवश्यकता है। डाटा फाइल तथा डाटा टेबलों सहित डाटा स्रोत/लेखापरीक्षा योग्य सत्व से प्राप्त डाटा सैटों का डाटा हस्तांतरण के तरीके सहित स्पष्ट रूप से उल्लेख किया जाना चाहिए। यदि आवश्यक हो, तो नोडल अधिकारियों की डाटा प्रबंधन करने हेतु पहचान की जानी चाहिए। डाटा सुरक्षा मामलों का पूर्ण सुरक्षा सुनिश्चित करने तथा डाटा सैटों तक किसी अनधिकृत पहुंच को रोकने के लिए पर्याप्त रूप से समाधान किया जाना चाहिए। डाटा की आवधिक रूप से उपलब्धता को सुनिश्चित करने के लिए लेखापरीक्षित सत्व के वरिष्ठ स्तरों पर पुष्टि प्राप्त करना आवश्यक होगा।
- 3.20 जैसे ही डाटा सैटों की संख्या में वृद्धि होगी, मॉडल हेतु डाटा प्रबंधन की जटिलता में वृद्धि होगी। बेहतर होगा कि मॉडल को सीधे लेखापरीक्षित सत्वों द्वारा उपलब्ध कराए गए रिस्टोर्ड डाटा पर ना बनाया जाए। इसके बजाय, मॉडल के सृजन हेतु संबंधित टेबलों का सार लिया जाए तथा उपयोग किया जाए। मॉडल तक पहुंच प्रयोक्ताओं को उनके पहुंच नियंत्रण प्रोफाइल जानने की आवश्यकता/उपयोग की आवश्यकता के आधार पर उपलब्ध कराई जानी चाहिए।

### **डाटा विश्लेषण प्रक्रिया का प्रलेखन**

- 3.21 विश्लेषण प्रक्रिया का प्रलेखन विश्लेषण परियोजना की योजना, निष्पादन तथा पर्यवेक्षण को सरल बनाता है। प्रलेखन प्रक्रिया के दौरान डाटा समेकता का रख-रखाव सहित विश्लेषण प्रक्रिया की समीक्षा तथा डाटा प्रबंधन के लिए उचित लेखापरीक्षा ट्रेल

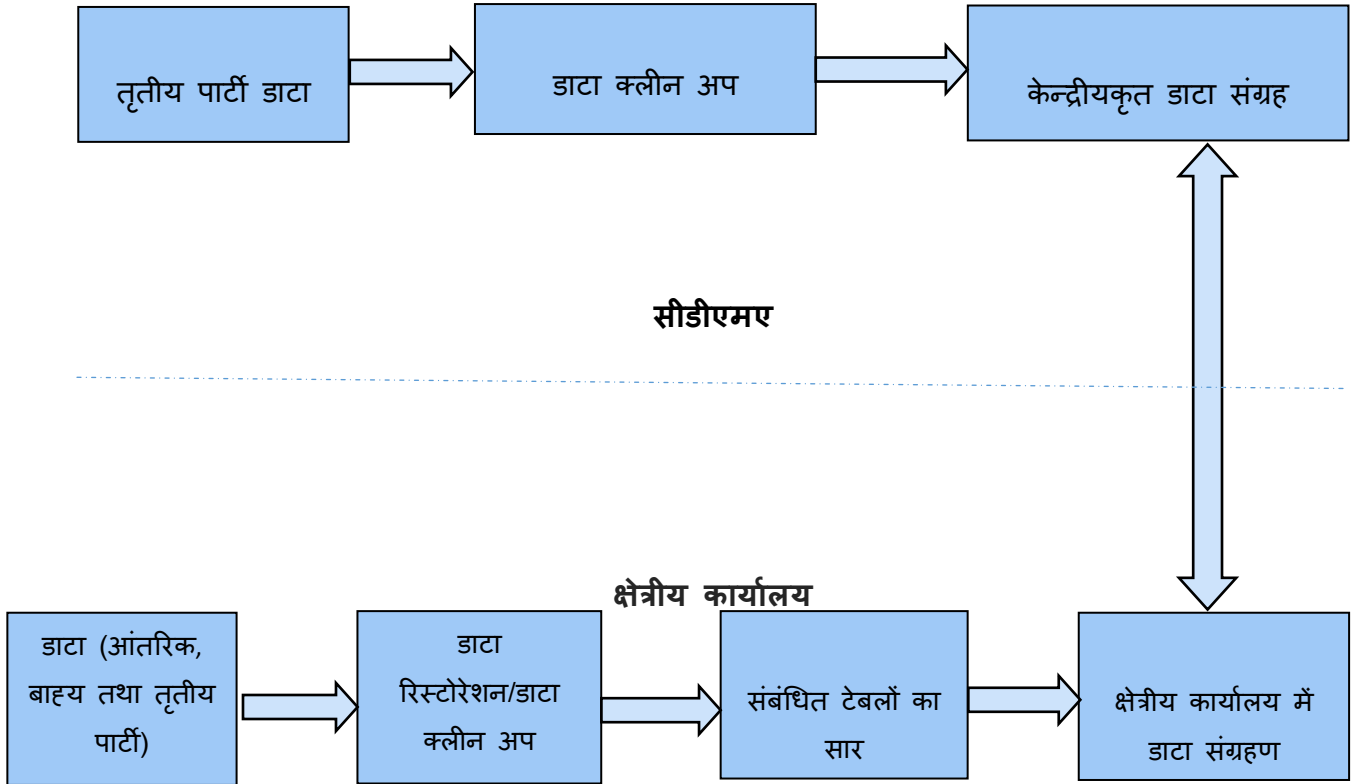
उपलब्ध कराने में भी सहायता करता है। लेखापरीक्षकों के परिणामों तथा निष्कर्षों के समर्थन के अलावा यह विश्लेषण प्रक्रिया को दोहराने में भावी लेखापरीक्षा दलों की सहायता करता है। डाटा विश्लेषण हेतु प्रलेखन में भारतीय लेखा एवं लेखापरीक्षा विभाग के लेखापरीक्षण मानदंडों का पालन किया जाना चाहिए। सभी प्रलेखन लेखापरीक्षक द्वारा हस्ताक्षरित तथा पर्यवेक्षी लेखापरीक्षा अधिकारी द्वारा प्रति हस्ताक्षरित होने चाहिए।

3.22 डाटा विश्लेषण कार्य के प्रलेखन में निम्नलिखित पहलू शामिल होने चाहिए:

- डाटा पहचान
- डाटा संग्रहण
- विश्लेषण सॉफ्टवेयर में डाटा इम्पोर्ट करना
- उपयोग की गई विश्लेषण तकनीक
- विश्लेषण के परिणाम
- डाटा विश्लेषण मॉडल
- लेखापरीक्षा में उपयोग से फीडबैक

### डाटा संग्रह

3.23 लेखापरीक्षा में साक्ष्य आधारित दृष्टिकोण यह अनिवार्य बनाता है कि लेखापरीक्षा उद्देश्यों की पहचान हेतु विभिन्न डाटा स्रोतों का उपयोग किया जाता है। जब डाटा विश्लेषण के माध्यम से जोखिम विश्लेषण लेखापरीक्षा प्रक्रिया का भाग बन जाता है तब यह आवश्यक है कि डाटा लेखापरीक्षा दल को शीघ्रता से उपलब्ध हो। इसे भारतीय लेखा एवं लेखापरीक्षा विभाग में प्रणालीगत डाटा संग्रहण तथा प्रबंधन प्रणाली के माध्यम से प्राप्त किया जा सकता है जिससे अंततः भारतीय लेखा एवं लेखापरीक्षा विभाग के लिए डाटा संग्रहण तैयार हो जाएगा। ऐसे डाटा संग्रह की परिकल्पना केंद्रीय स्तर तथा प्रत्येक क्षेत्रीय कार्यालय स्तर दोनों पर की जाती है। एक योजनाबद्ध डायग्राम नीचे दिया गया है:



चित्र 6 – डाटा संग्रह

### क्षेत्रीय कार्यालयों में डाटा संग्रह

3.24 क्षेत्रीय लेखापरीक्षा कार्यालय अपने कार्य क्षेत्र में आने वाले डाटा की पहचान तथा संग्रहण के लिए उत्तरदायी हैं। इसमें उनके प्रचालन क्षेत्र के विशेष आंतरिक डाटा स्रोत, लेखापरीक्षित सत्वों से प्राप्त डाटा तथा थर्ड पार्टी डाटा शामिल होंगे।

3.25 संग्रह बनाने के लिए डाटा संग्रहण एकबारगी प्रयास नहीं हैं। इसके बजाय यह एक समयावधि में निरंतर चलने वाली प्रक्रिया होगी जिसमें डाटा संग्रह करने के लिए डाटा की प्रणालीगत रूप से पहचान, संग्रहण, तैयारी, व्यवस्था, प्रबंधित तथा स्टोर किया जाएगा। प्रत्येक क्षेत्रीय लेखापरीक्षा कार्यालय में उनके कार्यक्षेत्र में डाटा प्रबंधन हेतु डाटा संग्रह करने के लिए निम्नलिखित उपायों का पालन किया जाना है;

- **डाटा पहचान** – डाटा प्रबंधन प्रणाली में पहला उपाय परिवेश में उपलब्ध डाटा स्रोतों की पहचान करना है। सभी क्षेत्रीय कार्यालयों को उनके कार्यक्षेत्र में उपलब्ध डाटा स्रोतों की पहचान करनी चाहिए। डाटा पहचान एक निरंतर चलने वाली प्रक्रिया है। लेखापरीक्षित सत्वों से प्राप्त डाटा के अलावा क्षेत्रीय कार्यालयों को उनके कार्यक्षेत्र से संबंधित थर्ड पार्टी डाटा की भी पहचान करनी चाहिए। विभिन्न डाटा स्रोतों से



संगृहीत डाटा सैटों को क्षेत्रीय कार्यालयों में अनुरक्षित डाटा संग्रह में भेजा जाना चाहिए।

- **डाटा मैपिंग** – एक बार डाटा स्रोतों की पहचान होने के बाद डाटा की क्षेत्रीय आधार पर मैपिंग की जानी चाहिए। यद्यपि डाटा स्रोतों को मुख्यतः विशेष क्षेत्र में प्रयोग हेतु डिजाइन किया जाय फिर भी इसे बहुल-क्षेत्रों में उपयोग किया जा सकता है। सभी डाटा स्रोतों के उपयोग का मास्टर डाटा निम्नलिखित फार्मेट में तैयार किया जाना चाहिए।

*डाटा उपयोग मास्टर टेबल प्रोफार्मा*

डाटा स्रोत का नाम	क्षेत्र जिससे यह संबंधित है प्राथमिक क्षेत्र	अन्य क्षेत्र जहां इसे उपयोग किया जा सकता है
माईनिंग एमआईएस डाटा	माइनिंग	वाणिज्यिक कर, परिवहन
यूडीआईएसई डाटा	शिक्षा	
जनगणना डाटा		शिक्षा, स्वास्थ्य आदि

चूंकि संगृहीत डाटा को डाटा विश्लेषण में उपयोग किया जाता है, अतः इसके माध्यम से प्राप्त जानकारी का डाटा उपयोग मास्टर टेबल में अद्यतित करने में उपयोग किया जाना चाहिए ताकि विभिन्न डाटा स्रोतों के बीच लिंकेज स्थापित किया जा सके।

- **डाटा तैयारी** – संग्रह में डाटा सैटों की स्टोरिंग से पूर्व डाटा संग्रह में इष्टतम स्टोरेज के लिए इन्हें डाटा तैयारी चरणों से गुजारना चाहिए।
- **डाटा अद्यतन**- क्षेत्रीय कार्यालयों को आवधिक रूप से डाटा सैटों को प्राप्त करने के लिए तंत्र की स्थापना करनी चाहिए। एक बार डाटा विश्लेषण मॉडलों के लिए अपेक्षित संबंधित डाटा सैट तैयार होने के बाद आगामी वर्षों में डाटा संग्रहण इन डाटा सैटों के लिए आवश्यक होगा जब तक कि उनकी संरचना में डाटा स्रोत का आशोधन नहीं हो जाता।
- **डाटा स्टोरेज** – चूंकि विभिन्न डाटा डम्प्स से प्राप्त संबंधित डाटा सैटों को डाटा संग्रह में डाला जाता है अतः संगृहीत डाटा डम्प्स को बाह्य स्टोरेज यंत्र में प्रणालीगत रूप से भंडारित किया जाना चाहिए।

- **मेटाडाटा** – डाटा स्रोतों, टेबल आदि के उचित मेटाडाटा को डाटा संग्रह के प्रबंधकों द्वारा अनुरक्षित किया जाना आवश्यक है। तीन इंटरलिंकड टेबलों के रूप में मेटाडाटा हेतु प्रारूप निम्नानुसार है:

#### डाटा स्रोतों की सूची

डाटा सैट का नाम	डाटा स्रोत का नाम	लेखापरीक्षित सत्व का नाम	क्षेत्र	निम्न से (तिथि/शुरुआती वर्ष)	निम्न तक (तिथि/वर्तमान वर्ष का उपलब्ध डाटा)	टेबलों की संख्या	डाटा आकार	टैग/की वर्ड <sup>20</sup>

#### डाटा स्रोत की टेबलों की सूची

डाटा स्रोतों का नाम	टेबल का नाम	विवरण	टेबल में कॉलमों की संख्या	पक्तियों की संख्या

#### प्रत्येक टेबल में फील्ड की सूची

डाटा स्रोतों का नाम	टेबल का नाम	फील्ड का नाम	फील्ड का प्रकार (सीएचएआर, आईएनटी), आदि	फील्ड का विवरण	अभ्युक्तियां

- 3.26 एक बार डाटा के तैयार तथा डाटा संग्रह में स्टोर हो जाने के बाद डाटा भविष्य में संदर्भ हेतु स्थायी रूप से स्टोर हो जाएगा। क्षेत्रीय कार्यालयों में डाटा विश्लेषण दल डाटा संग्रह के विकास तथा रख-रखाव में उपरोक्त सभी चरणों के लिए मुख्य रूप से उत्तरदायी होंगे।

#### केंद्रीय डाटा संग्रह

- 3.27 सीडीएमए डाटा हेतु एक डाटा संग्रह स्थापित करेगा जिसकी बहुल आईएण्डएडी कार्यालयों में प्रयोज्यता होगी। सीडीएमए आईएण्डएडी में उपयोग हेतु ऐसे डाटा सैटों का संग्रहण करेगा। क्षेत्रीय कार्यालय, परिभाषित अभिगम अधिकारों तथा प्रोटोकॉल के

<sup>20</sup> की वर्ड जिससे सर्चिंग करते समय डाटा सैटों की पहचान या इन्हे संदर्भित किया जा सकता है।

अनुसार केंद्रीय डाटा संग्रह में उपलब्ध डाटा सैटों तक पहुंचने में सक्षम होंगे। चयनित डाटा विश्लेषण मॉडल को भी केंद्रीय डाटा संग्रह में शामिल किया जाएगा।

### निरंतरता तथा उपलब्धता सुनिश्चित करना

- 3.28 किसी कार्यालय में डाटा विश्लेषण कार्यकलापों की निरंतरता को भारतीय लेखा एवं लेखापरीक्षा विभाग हेतु सूचना प्रणाली सुरक्षा हस्तपुस्तिका (दिसम्बर 2003)<sup>21</sup> में दिए गए कारबार निरंतरता प्रबंधन सिद्धांतों का पालन करके सुनिश्चित किया जाना चाहिए।
- 3.29 किसी कार्यालय में, डाटा तथा डाटा विश्लेषण परिणामों/मॉडलों की उपलब्धता<sup>22</sup> को पर्याप्त सुरक्षा की उपलब्धता के अलावा पर्याप्त नामसूची तथा वर्जन कंट्रोल के माध्यम से सुनिश्चित किया जाना चाहिए।

## 4. लेखापरीक्षा में डाटा विश्लेषण का उपयोग

- 4.1 डाटा विश्लेषण परिणामों का लेखापरीक्षा प्रक्रिया के दौरान इनसाइट या साक्ष्य प्राप्त करने हेतु लेखापरीक्षा<sup>23</sup> के किसी भी चरण पर उपयोग किया जा सकता है चाहे वह आयोजना, कार्यान्वयन या रिपोर्टिंग हो। लेखापरीक्षा योजना चरण पर मामलों की पहचान, यूनिट की योजना तथा नमूना डिजाइन डाटा विश्लेषण परिणामों से प्राप्त किया जा सकता है। लेखापरीक्षा कार्यान्वयन चरण पर लेखापरीक्षा विश्लेषण परिणामों से अपवादों, विचलनों की पहचान या मौजूदा स्थिति को भी चित्रित किया जा सकता है जिसे लेखापरीक्षा साक्ष्य के रूप में उपयोग किया जा सकता है। लेखापरीक्षा रिपोर्टिंग चरण पर, कार्यान्वयन चरण से प्राप्त डाटा विश्लेषण परिणामों को लेखापरीक्षा निष्कर्षों के बेहतर मूल्यांकन हेतु प्रस्तुत किया जा सकता है।
- 4.2 लेखापरीक्षण मानदंड प्रावधान करते हैं कि लेखापरीक्षक लेखापरीक्षा प्रक्रियाएं निष्पादित करेंगे जो कि लेखापरीक्षा रिपोर्ट के समर्थन हेतु पर्याप्त तथा उचित लेखापरीक्षा साक्ष्य उपलब्ध कराएंगी और यह साक्ष्य जानकार व्यक्ति को विश्वास

<sup>21</sup> भारतीय लेखा एवं लेखापरीक्षा विभाग के लिए सूचना प्रणाली सुरक्षा हस्तपुस्तिका (दिसम्बर 2003) का पैरा 8, भाग II (कार्यक्षेत्र विशिष्ट सुरक्षा अनुदेश)

<sup>22</sup> उपलब्धता का तात्पर्य अपेक्षित तरीके में, भारतीय लेखा एवं लेखापरीक्षा विभाग के लिए सूचना प्रणाली सुरक्षा हस्तपुस्तिका (दिसम्बर 2003) का भाग I (आईटी सुरक्षा नीति) समयबद्ध आधार पर पहुंच योग्य तथा उपयोग योग्य डाटा, सूचना तथा सूचना प्रणालियों की विशेषताओं से है।

<sup>23</sup> निष्पादन लेखापरीक्षा/अनुपालन लेखापरीक्षा/वित्तीय लेखापरीक्षा

दिलाने के लिए पर्याप्त (मात्रा) है कि निष्कर्ष तर्कसंगत है तथा उचित (गुणवत्ता) अर्थात् प्रासंगिक, वैध तथा विश्वसनीय (लेखापरीक्षण मानदंड, 2017 का पैरा 2.5.2.2 (क)) दोनों है। लेखापरीक्षा मानदंड यह भी अधिदेश देते हैं कि लेखापरीक्षक लेखापरीक्षा साक्ष्य का मूल्यांकन करेंगे तथा निष्कर्ष निकालेंगे। डाटा विश्लेषण परिणामों का लेखापरीक्षा साक्ष्य के रूप में तभी उपयोग किया जाना चाहिए जब वह लेखापरीक्षण मानदंडों की अपेक्षाओं का पालन करते हो।

- 4.3 विशिष्ट लेखापरीक्षा प्रक्रियाओं, जहां डाटा विश्लेषण परिणामों की निष्पादन, अनुपालन तथा वित्तीय लेखापरीक्षाएं करते समय काम में लाया जा सकता है, को अनुबंध 4 में संक्षेप में दिया गया है। तथापि, डाटा विश्लेषण परिणामों के उपयोग की मात्रा डाटा की उपलब्धता तथा डाटा विश्लेषण तकनीकों के उपयोग में क्षेत्रीय कार्यालय की परिपक्वता पर निर्भर करेगी।

#### विश्लेषण के लिए डाटा अधिग्रहण

- 4.4 भारतीय लेखा एवं लेखापरीक्षा विभाग (वित्तीय, अनुपालन, निष्पादन लेखापरीक्षाएं) द्वारा निष्पादित लेखापरीक्षाओं में डाटा विश्लेषण को काम में लाने हेतु प्रथम उपाय विश्लेषण हेतु प्रासंगिक डाटा की पहचान, संग्रहण तथा तैयारी करना है। लेखापरीक्षक को पृथक विभागों/सत्वों/क्षेत्रों की आरंभिक लेखापरीक्षा योजना<sup>24</sup> को अंतिम रूप देने से पूर्व सभी संबंधित डाटा सैटों आंतरिक, बाह्य तथा थर्ड पार्टी की पहचान करनी चाहिए। विभिन्न स्रोतों से प्राप्त इन डाटा सैटों को लिंक एवं इनका विश्लेषण किया जाना है जिसके परिणामस्वरूप लेखापरीक्षा के विभिन्न स्तरों पर इनका उपयोग किया जा सके।
- 4.5 यह संभव है कि सभी संबंधित डाटा सैटों की लेखापरीक्षा आरंभ होने से पूर्व पहचान न की गई हो। ऐसे मामले में, लेखापरीक्षक को नई डाटा उपलब्धता के बारे में सचेत रहना चाहिए। जब भी नए डाटा सैटों की पहचान एवं ऐसे डाटा तक एक्सेस किया जाता है तो जोखिम क्षेत्रों, हित क्षेत्रों, अपवादों या विचलनों की पहचान करने हेतु विश्लेषण किया जाना चाहिए।

---

<sup>24</sup> सीएजी द्वारा जारी वित्तीय प्रमाणन लेखापरीक्षण दिशानिर्देश, अनुपालन लेखापरीक्षण दिशानिर्देश तथा निष्पादन लेखापरीक्षण दिशानिर्देशों को संबंधित लेखापरीक्षाओं में यूनिटों के चयन तथा नमूना चयन संदर्भित किया जाए।

- 4.6 विभिन्न आधारभूत अवधारणाओं के साथ सांख्यिकीय जांच की जाती है। उसी समय डाटा सांख्यिकीय महत्व की विभिन्न विशेषताओं को स्पष्ट करता है। मान्य व्याख्या प्राप्त करने के उद्देश्य से प्रत्येक तकनीक/जांच के डाटा और पूर्वानुमानों या सीमाओं को समझने के लिए यह अनिवार्य है। अतः, लेखापरीक्षा निष्कर्षों के लिए यदि लेखापरीक्षक जांच परिणामों का इस्तेमाल करना चाहता हो तब निर्दिष्ट सांख्यिकीय जांचों को लागू करते हुए व्याख्या की वैधता प्रधान सांख्यिकीय अधिकारी या सांख्यिकीय सलाहकार द्वारा की जानी चाहिए।

### **लेखापरीक्षा योजना में डाटा विश्लेषणों का उपयोग**

- 4.7 आईएएण्डएडी में अधिकारी लेखापरीक्षा योजना के लिए जोखिम आधारित दृष्टिकोण अपनाते हैं। डाटा विश्लेषण साक्ष्य आधारित लेखापरीक्षा योजना की सहायता करता है और लेखापरीक्षा विषयवस्तु के संबंध में लेखापरीक्षा संसार तथा जोखिम क्षेत्रों को पहचानने में सहायता करता है। जबकि यह आईएएण्डएडी में मौजूदा जोखिम निर्धारण प्रक्रिया का स्थान ग्रहण नहीं करता है परन्तु उन्हें विशेष रूप से मजबूत करने के लिए डाटा विश्लेषण के पास यह संभावना है। तथापि, डाटा विश्लेषणों के परिणामों पर की जाने वाली निर्भरता की सीमा लेखापरीक्षक के निर्णय का मामला है।

### **वार्षिक लेखापरीक्षा योजना**

- 4.8 किसी कार्यालय में प्राप्त डाटा विश्लेषण क्षमता के साथ, यह परिकल्पित किया जाता है कि संगत आन्तरिक, बाह्य और थर्ड पार्टी डाटासेट्स और ऐनालेटिक मॉडल्स का डाटा संग्रहण जोखिम विश्लेषण प्रक्रिया में भरण करेगा। डाटा विश्लेषण परिणाम लेखापरीक्षा क्षेत्राधिकार के भीतर जोखिमों के सर्वांगी निर्धारण को उत्पन्न कर सकता है जिसका अन्य जोखिम निर्धारण मापदण्डों के अलावा वार्षिक लेखापरीक्षा योजना की तैयारी में उपयोग किए जाने चाहिए। वित्तीय डाटा पर आधारित डाटा विश्लेषण मॉडल है (जैसे वीएलसी डाटा, पीएफएमएस, इत्यादि के इस्तेमाल पर व्यवसाय प्रतिभा मॉडल्स) या अन्य क्षेत्र विशेष मॉडल्स इस कार्य में सहायता करेंगे और वार्षिक लेखापरीक्षा योजना की प्रक्रिया के लिए इनपुट उपलब्ध करवाएंगे।

### **योजना विशेष लेखापरीक्षाएं**

- 4.9 लेखापरीक्षाओं की ईकाई स्तर योजना के साथ साथ मूलभूत लेखापरीक्षा जांचों को करने के लिए साक्ष्य आधारित नमुना आरेख, लेखापरीक्षा उद्देश्यों के स्थापन सहित

विशेष लेखापरीक्षा के लिए जोखिम विश्लेषण और पहचान के मामलों में डाटा विश्लेषण को लगाया जा सकता है।

#### **जोखिम क्षेत्रों की पहचान करना जिसके कारण लेखापरीक्षा उद्देश्यों का स्थापन हुआ**

- 4.10 लेखापरीक्षा आरंभ करने से पहले निष्पादन लेखापरीक्षा दिशानिर्देश और अनुपालन लेखापरीक्षा दिशानिर्देश स्वत्व को समझने की आवश्यकता का उल्लेख करते हैं। वित्तीय लेखापरीक्षा में भी लेखापरीक्षा आरंभ करने से पहले व्यवसाय प्रक्रिया एवं प्रणाली की पहचान करने के लिए यह महत्वपूर्ण है। लेन-देन जहां इलेक्ट्रॉनिकली रिकार्ड किया गया है, वहां किसी भी प्रकार के लेखापरीक्षा (निष्पादन, अनुपालन, वित्तीय) के लिए बहु डाटा सेटस और आधारभूत संबंधित क्षेत्रों की पहचान का अधिकार देने से डाटा विश्लेषण की सभी जानकारी को सुविधाजनक बनाता है। यह व्यापक लेखापरीक्षा उद्देश्यों, उप उद्देश्यों तथा लेखापरीक्षा डिजाइन मैट्रिक्स की फ्रेमिंग को स्थापित करने में सहायता करेगा।

#### **वास्तविक जांचों के लिए नमूना ईकाईयों की पहचान करना**

- 4.11 डाटा विश्लेषण डाटा संबंधों, महत्वपूर्ण लेन-देन और बाहरी व्यक्तियों को इंगित करते हुए लेखापरीक्षित सत्त्वों के भीतर जोखिम क्षेत्रों की पहचान को सुगम बनाता है जिसके द्वारा वास्तविक जांचों के लिए लेखापरीक्षा इकाईयों के नमूनों के चयन के लिए अधिक वैज्ञानिक और केन्द्रित दृष्टिकोण प्रदान करता है। विभिन्न लेखापरीक्षा ईकाईयों के लिए मिश्रित जोखिम सूची पर पहुंचने और विभिन्न जोखिम मापदण्डों के भारित औसत अकों के आधार पर उन्हें रैंक देना लेखापरीक्षा ईकाईयों के नमूनों के चयन का निर्णय करने और वास्तविक जांचों की प्रकृति सीमा और समय निर्धारित करने की एक उद्देश्य विधि है।

#### **इकाई स्तर योजना**

- 4.12 लेखापरीक्षा इकाईयों के चयनित नमूनों में वास्तविक जांचों को लागू करने के लिए इकाई स्तर की योजना विशिष्ट लेन देन की पहचान करने का उल्लेख करती है। डाटा विश्लेषण दृष्टिकोण नमूना ईकाई के भीतर विशिष्ट मापदण्डों से विचलन की पहचान करने पर ध्यान केन्द्रित करेंगे। भ्रमणशील पार्टियों के लिए डाटा विश्लेषण परिणामों वाले ड्रिल डाऊन और फिल्टरिंग क्षमता युक्त गतिशील डेसबोर्ड विकसित किए जाएं।

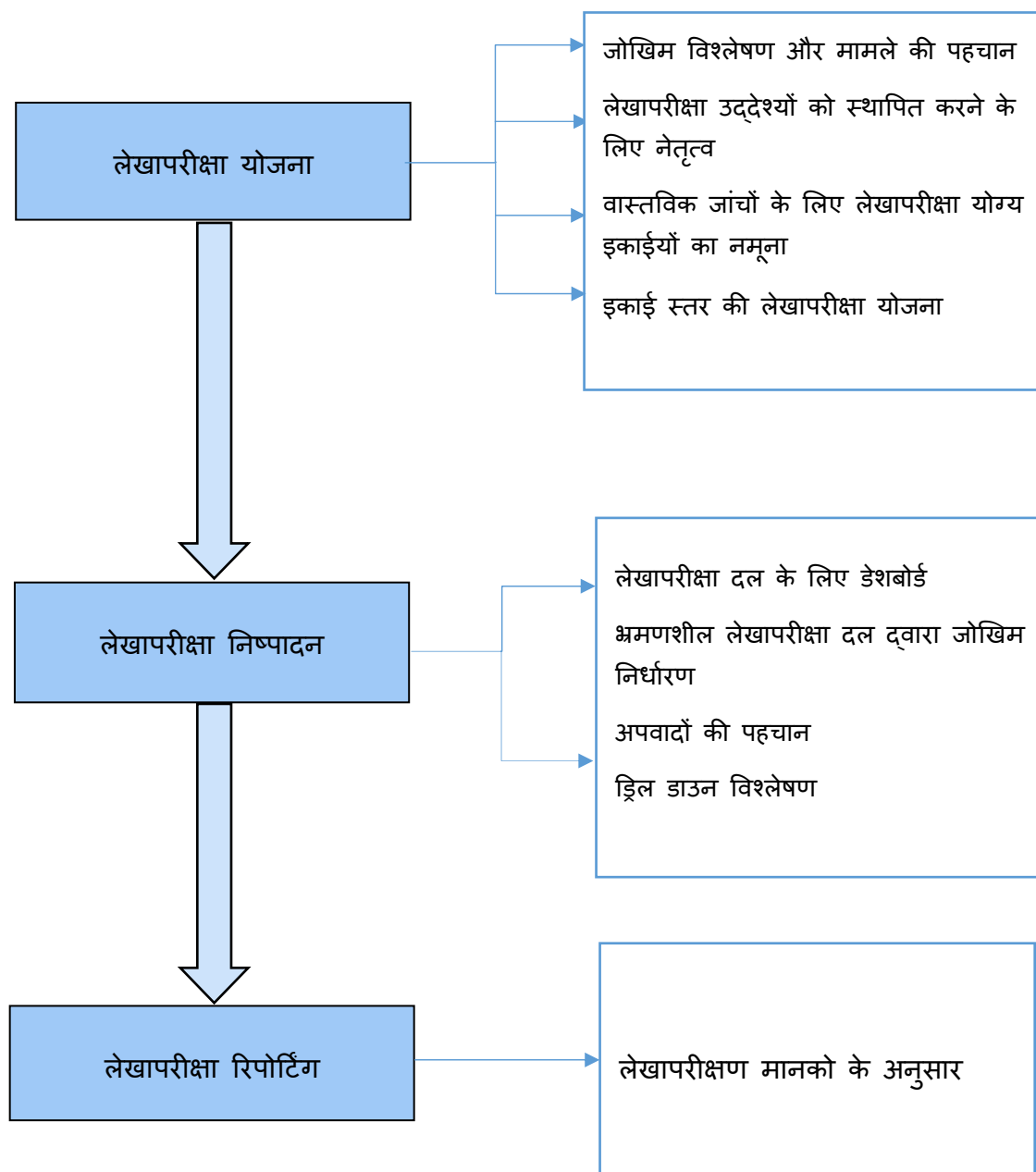
## लेखापरीक्षा निष्पादन में डाटा विश्लेषण का उपयोग

4.13 लेखापरीक्षा निष्पादन चरण के दौरान, लेखापरीक्षा योजना चरण के दौरान पहचाने गये लेखापरीक्षा अभिकथनों को सिद्ध करने के लिए साक्ष्य एकत्रित किया जाता है। डेटा विश्लेषण का उपयोग निम्नलिखित तौर-तरीकों में लेखापरीक्षा निष्पादन चरण के दौरान किया जा सकता है।

- लेखापरीक्षा योजना चरण के दौरान विकसित किये गये डैशबोर्ड/विश्लेषण परिणामों को भ्रमणशील लेखापरीक्षा दलों को प्रदान किया जाना चाहिए। इन डैशबोर्डों से लेखापरीक्षा योजना और लेखापरीक्षा इकाई चयन के विकास के कारण अधिक जानकारी के अलावा लेखापरीक्षा के लिए चयनित इकाई के बारे में अधिक जानकारी प्रदान करने में सहायता मिलेगी, डैशबोर्ड में लेन देन की सूची भी समाहित की जा सकती है जो लेखापरीक्षा नमूनों के चयन को प्रेरित कर सकती है। इस चरण पर लेखापरीक्षा के लिए चयनित यूनिट के विशिष्ट मुद्दों और उत्पन्न अपवाद रिपोर्टों को भी पहचाना जा सकता है। मुख्य विचलनों की पहचान करने के लिए समान ईकाइयों की तुलना में इसकी स्थिति के साथ-साथ लेखापरीक्षा के अन्तर्गत इकाई के जोखिम प्रतिमानों का मूल्यांकन करने के लिए सभी जानकारियां लेखापरीक्षकों को सुसज्जित करती हैं। यह इकाई के लिए विशेष नमूना चयन पर ध्यान केंद्रित करने में लेखापरीक्षक की मदद करेगा यदि लेखापरीक्षा योजना स्तर पर पहले ही नहीं किया गया हो।
- भ्रमणशील लेखापरीक्षा दल लेखापरीक्षाओं में डाटा विश्लेषण के ज्ञान को लागू करने के लिए सक्षम हो जाएंगे, जब वे लेखापरीक्षाओं के दौरान इलेक्ट्रॉनिक डाटा का उपयोग करेंगे। इस प्रकार, लेखापरीक्षा दल अपने लेखापरीक्षा दावों को सिद्ध करने के लिए ड्रिल डाउन तकनीक का उपयोग कर सकते हैं। उन्हें लेखापरीक्षा प्रक्रिया के दौरान प्राप्त अधीनस्थ डाटा सेट के लिए भी डाटा विश्लेषण को लागू करना चाहिए (यदि इलेक्ट्रॉनिकली उपलब्ध हो) जो कॉरपोरेट या संगठन स्तर डाटा के आवश्यक रूप से भाग नहीं बनते हैं।

## लेखापरीक्षा रिपोर्टिंग में डाटा विश्लेषण का उपयोग

- 4.14 लेखापरीक्षा प्रक्रिया पणधारियों के लिए लेखापरीक्षा के परिणामों का संप्रेषण करने के लिए एक रिपोर्ट तैयार करना शामिल करती है जिसका लेखापरीक्षण मानकों में परिकल्पित रिपोर्टिंग आवश्यकताओं के साथ अनुपालन करना चाहिए।
- 4.15 लेखापरीक्षा योजना निष्पादन और रिपोर्टिंग में डाटा विश्लेषणों के प्रमुख उपयोगों की एक योजनाबद्ध प्रस्तुती नीचे दी गई है।



चित्र 7- लेखापरीक्षा प्रक्रिया के विभिन्न चरणों में डाटा विश्लेषण का उपयोग



*अनुबंध*

## अनुबन्ध 1 (संदर्भ पैरा 1.10)

### डाटा विश्लेषण कार्य हेतु भूमिकाएं और उत्तरदायित्व

#### सीडीएमए की भूमिकाएं और उत्तरदायित्व

आईएण्डएडी में डाटा विश्लेषण के सम्पूर्ण उपयोग के लिए सीडीएमए एक सलाहकार और सहायक की भूमिका निभाएगा। सीडीएमए क्षमता निर्माण, केन्द्रीय स्तर पर थर्ड पार्टी डाटा एकत्रित करना, नये साफ्टवेयर की पहचान करना, विभिन्न विश्लेषण तकनीकों/विश्लेषणा प्रतिरूपों की प्रयोज्यता का निर्धारण करना और उन्हें आईएण्डएडी में प्रचारित करने के माध्यम से सरल बनाएगा। सीडीएमए जहां भी आवश्यक हो क्षेत्रीय कार्यालयों को उनके डाटा विश्लेषण प्रयासों में तकनीकी सहायता उपलब्ध करवाएगा।

मुख्यालय में कार्यात्मक स्क्व के साथ परामर्श से डाटा विश्लेषण नमूनों की संवीक्षा एवं अनुमोदन सीडीएमए द्वारा किया जाएगा

#### क्षेत्रीय कार्यालयों में भूमिकाएं और उत्तरदायित्व

डाटा विश्लेषण परियोजना क्षेत्रीय कार्यालय के भीतर कार्यात्मक समूह की जिम्मेदारी है। डाटा विश्लेषण समूह परियोजना को आवश्यक तकनीकी सहायता प्रदान करेगा।

डाटा विश्लेषण क्रियाकलाप	डाटा विश्लेषण समूह	क्षेत्रीय कार्यालयों में कार्यात्मक समूह
1. डाटा की पहचान		प्राथमिक उत्तरदायित्व
2. डाटा संग्रहण	डाटा संग्रहण प्रक्रिया में तकनीकी सहायता	प्राथमिक उत्तरदायित्व
3. डाटा पुनः स्थापन	प्राथमिक उत्तरदायित्व	
4. डाटा की तैयारी	प्राथमिक उत्तरदायित्व	तैयारी में डोमैन विशेषज्ञता उपलब्ध कराना
5. डाटा विश्लेषण और मॉडल का निर्माण	तकनीकी सहायता	प्राथमिक उत्तरदायित्व
6. मॉडल की तैयारी	प्राथमिक उत्तरदायित्व	
7. मॉडल हेतु डाटा का आवधिक संग्रहण	तकनीकी सहायता	प्राथमिक उत्तरदायित्व
8. डाटा जमाकर्ता और प्रबंधन	प्राथमिक उत्तरदायित्व	
9. दस्तावेजीकरण	जहां डाटा विश्लेषण समूह प्राथमिक जिम्मेदारी रखते हैं वहां उनसे ऊपर क्षेत्रों पर दस्तावेज प्रदान करना दस्तावेजीकरण का संकलन करना	जहां कार्यात्मक समूह प्राथमिक जिम्मेदारी रखते हैं वहां उनसे ऊपर क्षेत्रों पर दस्तावेज प्रदान करना
10. डाटा विश्लेषण हेतु बुनियादी ढांचा	तकनीकी आवश्यकताओं की पहचान करना	प्राथमिक जिम्मेदारी प्रशासनिक स्क्व की होगी।

## अनुबंध 2 (संदर्भ पैरा 2.13)

डाटा की पूर्णता, अनुरूपता और सत्यनिष्ठा के प्रमाणपत्र का सांकेतिक टेम्पलेट (डाटा प्राप्त करते समय लेखापरीक्षित स्वत्व से संग्रहीत किया जाना है)

आज (तारीख) को <XXXXX> (डाटा देने वाले नॉडल अधिकारी का नाम एवं पद) द्वारा <XXXXX> (बाहरी उपकरण के मामले में) के रूप में यथावत चिन्हित तथा हस्ताक्षरित/प्राधिकृत किसी बाह्य भंडार उपकरण/आनलाईन प्रदत्त पर मंत्रालय/ विभाग/ \_\_\_\_\_ (डाटा देने वाले स्वत्व का नाम) द्वारा अनुरक्षित \_\_\_\_\_ से \_\_\_\_\_ तक की अवधि के लिए \_\_\_\_\_ (डाटा बेस का नाम) के संबंध में कार्यालय \_\_\_\_\_ (लेखापरीक्षा कार्यालय का नाम) को डाटा डम्प दिया गया।

यह सत्यापित किया जाता है कि:-

- (i) लेखापरीक्षा के साथ इस डाटा को साझा करने के लिए \_\_\_\_\_ (लेखापरीक्षित स्वत्व का नाम) द्वारा अधिकारी प्राधिकृत हैं और वे सूचना प्रादुर्गिकी अधिनियम 2008 के सुसंगत प्रावधान को समझते हैं।
- (ii) डाटा डम्प पूरी तरह पूर्ण और संपूर्ण वास्तविक डाटा है।
- (iii) मूल डाटा में कोई विलोपन, हेर-फेर या उपरिलेखन नहीं है।
- (iv) कोई डाटा अननुरूपता नहीं है और एक प्रणाली से दूसरी प्रणाली में डाटा स्थानांतरण के दौरान या बेकअप या चोरी/ड्रैकिंग के कारण डाटा हानि नहीं हुई थी।
- (v) डाटा की कोई क्षति अर्थात: किसी भी तरह से कंप्यूटर संसाधनों के नष्ट, परिवर्तन, रूपांतरण, लोप या पुनः व्यवस्थित द्वारा नहीं हुई है।

मुख्य मापदंडों पर सारांश सूचना- डाटा की पूर्णता, अनुरूपता और सत्यनिष्ठा को सुनिश्चित करने के लिए लेन-देन की कुल संख्या, प्रथम तथा अन्तिम लेन-देन की तारीख एवं विवरण और विभिन्न संख्या डाटा क्षेत्रों का द्रुतान्वेषण कुल भी दिया गया है।

(नाम, पदनाम, ई-मेल एवं प्राधिकृत अधिकारी के हस्ताक्षर)

दिनांक

स्थान



अनुबन्ध 3 (संदर्भ पैरा 3.13)

**डाटा विश्लेषण परिणामों से प्राप्त होने वाली पूरी जानकारी**

सभी क्षेत्रों के लिए वर्तमान कर मांग एवं वर्तमान कर संग्रहण के संबंध में डाटा पर सांख्यिकीय विश्लेषण किया गया है। प्रत्येक क्षेत्र के लिए मांग एवं संग्रह डाटा का सहसंबंध परिकलित किया गया है। सूची बनाने तथा प्रलेख सांख्यिकीय निष्कर्षों (सहसंबंध विश्लेषण) के लिए टेम्पलेट तथा इससे पूरी जानकारी निकालना नीचे दर्शाया गया है।

संदर्भ आंकड़े /तालिका (कल्पना आकृति अथवा सांख्यिकीय परिणाम दिखाने वाली तालिका)	सांख्यिकीय निष्कर्ष (सांख्यिकीय परिणाम)	सृजित पूरी जानकारी (सांख्यिकीय निष्कर्ष की व्याख्या)	लेखापरीक्षा के लिए केन्द्रित क्षेत्र	क्या मॉडल में शामिल किए जाने के लिए काई पूरी जानकारी है, यदि कोई हो
वर्तमान मांग तथा वर्तमान संग्रहण के बीच सहसंबंध दिखाने वाली जोन वार तालिका	जोन XX तथा YY में 0.533, 0.421 के आंकड़े वर्तमान मांग तथा वर्तमान संग्रहण के बीच खराब सहसंबंध दर्शाते हैं।	वर्तमान मांग के प्रति कर संग्रहण संतोषजनक नहीं है।	वर्तमान मांग के प्रति ऐसे खराब संग्रहण के कारणों की खोज की जानी चाहिए।	✓

## अनुबन्ध 4 (संदर्भ : पैरा 4.3)

### विभिन्न प्रकार की लेखापरीक्षाओं में डाटा विश्लेषण

#### क. निष्पादन लेखापरीक्षा

निष्पादन लेखापरीक्षा दिशानिर्देश 2014 के अध्याय 4 के अनुसार, निष्पादन लेखापरीक्षाओं की वैयक्तिक योजना, सत्व/कार्यक्रम को समझना किसी वैयक्तिक निष्पादन लेखापरीक्षा का आरम्भिक बिन्दु है, जिसमें विभिन्न रूपों तथा स्रोतों में सूचना जैसे कि इलेक्ट्रॉनिक डाटा बेस, प्रबंधन सूचना प्रणाली, एमआईएस प्रतिवेदनों, वेबसाइट से सूचना आदि की समीक्षा शामिल है। अध्याय लेखापरीक्षा रूपरेखा तैयार करने वाले अभिगमों तथा प्रणालियों, विभिन्न विश्लेषणात्मक तकनीकों के उपयोग, लेखापरीक्षा रूपरेखा मैट्रिक्स तैयार करना एवं डाटा संग्रहण एवं विश्लेषण विधियां निर्धारित करना जिसमें डाटा विश्लेषण भी शामिल है, की भी परिकल्पना करता है। निम्नलिखित कदम निष्पादन लेखापरीक्षा में डाटा विश्लेषण के उपयोग हेतु अपनाई जाने वाली सामान्य कार्यप्रणाली का उल्लेख करते हैं।

- सभी संबंधित डाटा सैट्स की निष्पादन लेखापरीक्षा के योजना चरण पर पहचान की जानी चाहिए। डाटा सैटों पर फॉर्मेट, साईज, एक्सेस मोड, पर मूल सूचना लेखापरीक्षित सत्वों से एकत्रित करनी चाहिए। डाटा इन दिशानिर्देशों के अध्याय 2 में डाटा तैयारी हेतु निर्धारित प्रक्रियाओं के अनुसार आवश्यकतानुसार तैयार किया जाना चाहिए।
- डाटा विश्लेषण लेखापरीक्षित सत्वों के प्राथमिक डेटाबेस पर आरम्भ किया जाय। यह एमआईएस डेटाबेस, लेखापरीक्षित सत्वों के लेनदेन डेटाबेस हो सकता है। विश्लेषण डाटा वर्गीकृत करने, भौगोलिक/प्रशासनिक भिन्नताओं को समझने, समयाोपरि भिन्नताओं (प्रवृत्ति विश्लेषण) आदि के लिए विभिन्न दृश्य एवं विवरणात्मक सांख्यिकीय तकनीकों का उपयोग कर डाटा खोज के साथ आरम्भ किया जाना चाहिए। जीआईएस मानचित्रों का प्रयोग विभिन्न मापदंडों के स्थानिक वितरण को समझने में सहायता करेगा। उदाहरण के लिए यदि यह निष्पादन लेखापरीक्षा स्वास्थ्य क्षेत्र पर है तो विश्लेषण यह देखने के लिए किया जा सकता है
  - क्या स्वास्थ्य सूचक उपलब्ध हैं।
  - क्या स्वास्थ्य संरचना विश्लेषण को विभिन्न राज्यों/जिलों में सूचित किया जा रहा है।
  - विभिन्न क्षेत्रों तथा कई वर्षों से स्वास्थ्य सूचकों में विभिन्नता

- इस विश्लेषण से, लेखापरीक्षक सत्व/योजना की प्रकृति को समझ सकता है तथा रूचि के क्षेत्र की पहचान कर सकता है।
- निष्पादन लेखापरीक्षा मुख्यतः यह सत्यापित करना चाहती है कि क्या योजना/कार्यक्रम वांछित परिणामों की प्राप्ति में सफल है या दक्षता से कार्यान्वित किया जा रहा है। लेखापरीक्षक को विभिन्न सरकारी व्यवधानों को पहचानने की आवश्यकता है तथा फिर यह देखना कि यह परिणामों को कैसे प्रभावित कर रहा है। इसलिए अगला सोपान सरकार के परिणामों/आउटपुट तथा विभिन्न इनपुट कारकों के बीच के संबंध को समझना होगा। इसे डाटा के सलाईसिंग एवं डाइसिंग से तथा प्लॉटों के बिखराव, सहसंबंध, प्रतिगमन आदि जैसी विश्लेषण तकनीक उपयोग कर प्राप्त किया जा सकता है। उदाहरण के लिए यदि यह समीक्षा शिक्षा पर है तो विश्लेषण उन कारकों की पहचान करने के लिए किया जा सकता है जो नामांकन, खारिज करने वाले जैसे परिणामों में सुधार कर सकते हैं।
- इस चरण में, मल्टीपल डाटा स्रोत तथा डाटा सेट की जांच की जा सकती है जो प्राथमिक डाटा सेट से जुड़े हो सकते हैं। परिणामों के मापन के लिए, लेखापरीक्षक को केवल लेखापरीक्षित सत्वों द्वारा दिए गए आंकड़ों पर निर्भर रहने की आवश्यकता नहीं है। तीसरी पार्टी के विश्वसनीय डाटा सेट्स उपयोग किये जा सकते हैं। विषय वस्तु की बढ़ती संख्या का डाटा सेट (सैंकडरी डाटा सेट) तुरन्त प्राप्य है जो लेखापरीक्षकों को सुसंगत मुद्दों के आलोचनात्मक विश्लेषण के लिए अनुमत है तथा कुछ अप्रतिरोध्य प्रश्नों का समाधान करते हैं।
- इन पूर्ववर्ती चरणों सहित, लेखापरीक्षक संगठनों में जोखिमों/रूचियों को समझ सकता है जिसका लेखापरीक्षा में समाधान करना होगा। विभिन्न डाटा सेट का विश्लेषण न केवल समग्र परिप्रेक्ष्य बल्कि लेखापरीक्षा उद्देश्यों/उप उद्देश्य को परिभाषित करने के लिए साक्ष्य आधारित अभिगम प्रदान करता है। उसी समय, पर लेखापरीक्षक को अवगत होना चाहिए कि जोखिम मापदंड हो सकते हैं जिन्हें विद्यमान डाटा सेटों में अधिकृत नहीं किया गया है। इस प्रकार से पायलट स्टडी आदि से पूर्व सत्वों/सूचना को समझना साक्ष्य आधारित अभिगम को सशक्त करने के लिए उपयोग किया जाना चाहिए।
- एक बार लेखापरीक्षा उद्देश्य परिभाषित किए जाने पर अगला सोपान जोखिम बोध पर आधारित यथोचित जांचों के लिए नमूना यूनिटों की पहचान करना होगा। चूंकि मल्टीपल जोखिम सूचकों की डाटा विश्लेषण से पहचान होगी, इसलिए विभिन्न जोखिम मापदंडों के भारित औसत पर आधारित समग्र स्कोर किया जा सकता है तथा इस स्कोर पर सैंपलिंग की जा सकती है। एक भारित जोखिम स्कोर कारकों के लिए

भी नियत किया जा सकता है जिसे डाटा विश्लेषण टूल्स से पहचाना नहीं गया है एवं समग्र स्कोर में सम्मिलित किया नहीं गया है।

- भ्रमणकारी टीमों उनको दिए गए डेसबोर्ड की सहायता से यूनिट स्तरीय योजना को अब प्रारम्भ कर सकती है। आधुनिक डाटा विश्लेषण टूल्स में उपलब्ध डायनेमिक ड्रिल डाउन तथा फिल्टरिंग क्षमता से, लेखापरीक्षा टीमों प्रत्येक नमूना यूनिट के संबंध में जोखिमों की पहचान कर सकती है तथा उप यूनिट में लेखापरीक्षा योजना बना सकती है।
- रिपोर्टिंग चरण में, पणधारियों को लेखापरीक्षा निष्कर्षों को प्रस्तुत करने की गुणवत्ता विभिन्न दृश्य तकनीकों के माध्यम से सुधारी जा सकती है।

## ख. अनुपालन लेखापरीक्षा

अनुपालन लेखापरीक्षण दिशानिर्देशों के पैरा 2.25 के अनुसार, लेखापरीक्षा योग्य सत्त्वों की जोखिम प्रोफाइलिंग संगठन के भारी जोखिम क्षेत्रों/गतिविधियों की पहचान किए जाने की आवश्यकता है। दिशानिर्देश जोखिम क्षेत्रों को पहचानने के लिए बड़े डाटा का लाभ लेने और सामाजिक-आर्थिक सर्वेक्षण, बजट/वीएलसी तथा अन्य डाटा स्रोतों जैसे विभिन्न डाटा का उपयोग करने का उल्लेख करते हैं। इसी प्रकार अनुपालन लेखापरीक्षण दिशानिर्देश के अध्याय 4 के अनुसार, वैयक्तिक अनुपालन लेखापरीक्षा यूनिट की योजना बनाए जाने की आवश्यकता है।

अनुपालन लेखापरीक्षा में, लेखापरीक्षक जिन प्राथमिक प्रश्नों का सामना करता है उन मामलों की पहचान करना है जहां विधि/नियम की अनुपालना नहीं की गई है। लेखापरीक्षा योजना परिप्रेक्ष्य से यथोचित जांच के लिए यूनिटों की पहचान जहां ऐसे अननुपालन की सम्भावना है, निर्णायक होगी। डाटा विश्लेषणात्मक मॉडल्स इस काम में तथा ऐसे अननुपालन के मामलों पर व्यवस्थित ढंग से पहुँच सकने में सहायक हो सकते हैं।

विभिन्न तकनीकों का उपयोग करते हुए सुसंगत डाटा सैटों पर डाटा विश्लेषण पहचान करने तथा विभिन्न मानदंडों पर सभी यूनिटें वर्गीकृत करने में सहायक होगा। ये जोखिम तत्व क्षेत्र से क्षेत्र भिन्न होंगे। अभिगम क्षेत्र के संदर्भ में मल्टीपल जोखिम सूचक की पहचान होगी। जोखिम मानदंडों के कुछ उदाहरण हैं:

- व्यय
- गत वर्षों से व्यय में असाधारण भिन्नता
- विलम्ब (प्रचालनों के मामले में)
- प्राप्ति लेखापरीक्षा के मामले में आय/बिक्री अनुपात से कम कर



एक बार जोखिम मानदंड की पहचान हो जाने पर प्रत्येक मानदंड के लिए सत्वा की प्रत्येक लेखापरीक्षा यूनिटों के लिए जोखिम स्कोर निर्धारित किए जा सकते हैं। मल्टीपल मानदंडों के भारित औसत पर आधारित समग्र स्कोर की गणना की जा सकती है तथा यथोचित जांच के लिए सैंपल यूनिटों को इस स्कोर के आधार पर चयनित किया जा सकता है। सैंपलिंग का अगला स्तर यथोचित जांच के लिए इन यूनिटों में से प्रत्येक लेनदेन का चयन किया जाएगा। ऐसे भारी जोखिम लेन देनों की पहचान निम्न द्वारा की जा सकती है:

- नियम स्थिति को शामिल करना जिसके अनुपालन की प्रश्नों/डेशबोर्ड के माध्यम से मांग की जा रही है जैसे कि अपवाद/अनुपालन के मामलों की पहचान की जा सकती है।
- पैटर्न, क्लस्टर अथवा आउटलाईअर की पहचान करने के लिए स्केटर प्लॉट, बॉक्स प्लाट आदि जैसी दृश्य तकनीकें।

अधिकतर डाटा विश्लेषण टूल्स में उपलब्ध ड्रिल डाउन तथा फिल्टरिंग क्षमता के साथ, एक डेशबोर्ड यथोचित जांच के लिए चयनित सभी यूनिटों के लिए बन सकता है। डाटा की आवधिक उपलब्धता सुनिश्चित करके, विश्लेषण एक डाटा विश्लेषण मॉडल को आगे ले जाने के लिए कुछ वर्षों के लिए दोहराया जा सकता है। एक निष्पादन लेखापरीक्षा के लिए किए गए विश्लेषण को उपयुक्त संशोधनों के माध्यम से क्षेत्र में अनुपालन लेखापरीक्षा के लिए मूल मॉडल में परिवर्तित किया जा सकता है।

### ग. वित्तीय लेखापरीक्षा

वित्तीय विवरणों की लेखापरीक्षा का उद्देश्य वित्तीय विवरणों में अभिप्रेत उपयोगकर्ताओं के विश्वास को बढ़ाना है। इसे लेखापरीक्षक द्वारा एक मत की अभिव्यक्ति के माध्यम से प्राप्त किया जाता है कि क्या वित्तीय विवरण लागू वित्तीय रिपोर्टिंग फ्रेमवर्क के अनुसार, सभी सार्थक प्रकार से तैयार किए जाते हैं अथवा एक निष्पक्ष प्रस्तुतीकरण वित्तीय रिपोर्टिंग फ्रेमवर्क के अनुसार तैयार किए गए कि क्या वित्तीय विवरणों के मामले में क्या वित्तीय विवरण सभी सार्थक प्रकार से अथवा निष्पक्ष रूप से प्रस्तुत किए गए हैं, अथवा उस फ्रेमवर्क के अनुसार एक सत्य एवं निष्पक्ष राय देते हैं।

एक संगठन अथवा संघ सरकार/राज्य सरकार के वित्तीय विवरणों की लेखापरीक्षा में, वर्णनात्मक विश्लेषण तथा दृश्य चित्रण का उपयोग संव्यवहारों के विभिन्न वर्गों, लेखा शेषों, विशिष्ट अनुदानों अथवा प्रकटनों में पूरी जानकारी समझने तथा प्राप्त करने में लेखापरीक्षक की सहायता कर सकता है, जो अस्पष्ट भिन्नता अथवा असामान्यता की सम्भावना को प्रकट

करता है। यथोचित जांच के लिए नमूनों को डाटा विश्लेषण से आहरित सभी जानकारी के आधार पर आहरित किया जा सकता है।

वित्तीय लेखापरीक्षकों द्वारा की जाने वाली नियमित जांचे स्वचालित की जा सकती है तथा एक मॉडल में बनाई जा सकती है, जो आगामी वर्षों से संबंधित डाटा के संबंध में डाटा सहित अद्यतित की जा सकती है। वर्षों से वित्तीय व्यवहारों की प्रवृत्ति किसी असामान्य व्यवहार अथवा पैटर्न की पहचान करने में सहायक होगी। इस प्रकार तैयार डेशबोर्ड क्षेत्रीय लेखापरीक्षा करने में लेखापरीक्षक की सहायता करेगा।